Electronic Transactions on Numerical Analysis.
Volume 57, pp. 153–192, 2022.

ISSN 1068–9613.
DOI: 10.1553/etna_vol57s153

ETNA
Kent State University and
Johann Radon Institute (RICAM)

REGULARIZED FUNCTIONAL MATCHING PURSUIT FOR THE SPHERICAL MULTIPLE-SHELL ELECTRO-MAGNETOENCEPHALOGRAPHY PROBLEM*

SARAH LEWEKE† AND VOLKER MICHEL†

Abstract. Reconstruction of the neuronal current inside the human brain from non-invasive measurements of the magnetic flux density via magnetoencephalography (MEG) or of electric potential differences via electroencephalography (EEG) is an invaluable tool for neuroscientific research, as it provides measures of activity in the brain. However, it is also a severely ill-posed inverse problem. Assuming spherical geometries, we consider the spherical multiple-shell model for the inverse MEG and EEG problem and apply the regularized functional matching pursuit algorithm (RFMP) for its solution. We present a new convergence proof for the RFMP for operators between two infinite-dimensional Hilbert spaces. Moreover, we utilize the complementarity of EEG and MEG data to combine inversions of simultaneous electric and magnetic measurements. Finally, we test the algorithm numerically on synthetic data using several Sobolev norms as penalty term and apply it to real data.

Key words. electroencephalography, greedy algorithms, ill-posed problems, integral equation, inverse problems, magnetoencephalography, regularization methods, regularized functional matching pursuit, Sobolev spaces

AMS subject classifications. 42C10, 45B05, 46C07, 46N40, 47A52, 65R30, 65R32

1. Introduction.

In this paper, we apply the regularized functional matching pursuit algorithm (RFMP) to the inverse electro-magnetoencephalography problem in the spherical multiple-shell model. The RFMP was originally invented for the inversion of gravitational data and normal-mode anomalies; see [5, 15, 16]. Especially for the inversion of gravitational data, several different approaches with different function systems and bases exist; e.g., spherical harmonics, wavelets, Slepian functions, and splines. Since each of these systems has advantages but also disadvantages, an algorithm was sought that allows several different function types to be combined. This leads to a dictionary that may contain linearly dependent functions and may be redundant. The aim of this method is to acquire all the advantages of the different systems and to reduce their respective disadvantages.

In the meantime, several variants of the RFMP have been presented in the literature and applied to several ill-posed inverse problems in the geosciences; see [17, 18, 20, 44, 49, 50]. They are based on a greedy algorithm introduced in [41] and enhanced in [58] that is called the matching pursuit algorithm. The novelty of the RFMP is its applicability to ill-posed inverse problems.

The RFMP iteratively builds a linear combination of dictionary elements to approximate the minimizer of the Tikhonov functional. To this end, the dictionary element and the linear factor of the next iteration are chosen such that the Tikhonov functional is optimally reduced. An improvement of the method, called the regularized orthogonal functional matching pursuit (ROFMP), is described in [50, 57] and limits the search of the dictionary elements in each iteration to those in the orthogonal complement of the ones used in the previous iterations.

In [15, 16], convergence results of the RFMP are stated for the first time. Therein, only operators mapping from $\mathrm{L}^2(G,\mathbb{R})$ to $\mathbb{R}^\ell$ are considered, where the usual classes of $\mathrm{L}^2(G,\mathbb{R})$ denote the space of square-integrable functions $G \to \mathbb{R}$ with respect to the Lebesgue measure and $G \subset \mathbb{R}^d$ is a compact domain. The convergence results of the RFMP consist of two statements: the sequence of approximations produced by the RFMP converges, and the RFMP is a convergent regularization method. The convergence result was transferred to an arbitrary

*Received March 14, 2022. Accepted August 25, 2022. Published online on September 19, 2022. Recommended by Ronny Ramlau.

†Geomathematics Group, Department of Mathematics, University of Siegen, Walter-Flex-Straße 3, 57068 Siegen, Germany ({leweke,michel}@mathematik.uni-siegen.de).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

separable Hilbert space as the domain of the operator in [47, 57]. In this context, the statement concerning the convergence of the sequence of approximations produced by the RFMP was improved by the authors in [47].

In [33, 34], the RFMP was considered as a particular case of the newly introduced regularized weak functional matching pursuit (RWFMP), which is applicable to operators between two (infinite-dimensional) Hilbert spaces. This approach yields a characterization of the convergence order of the regularization induced by the RWFMP, which also follows for the RFMP. In this work, we present an alternative approach to achieve this convergence result without the technique required for the RWFMP. The transition to the infinite-dimensional setting in the operator range requires only a small adjustment to the proof stated in [47].

Finally, we apply the RFMP algorithm to the spherical multiple-shell electro-magneto-encephalography problem; see [38]. This problem deals with the reconstruction of the vector-valued neuronal current inside the human brain from non-invasive measurements of the electric potential on the scalp and the induced magnetic field outside the head. Instead of the real brain geometry, a spherical geometry is assumed, and multiple shells are employed to represent the different conductive tissues that the signals need to traverse. In [38], the problem is cast as Fredholm integral equations and shown to be severely ill-posed, since the singular values decay exponentially fast and the kernels of the integral operators are infinite-dimensional. Normally, joint inversions of different data types require sophisticated methodologies; see, e.g., [23]. However, we prove in this paper for this particular application that the simultaneous inversion does not yield more information than the combination of independently inverted magnetoencephalography (MEG) and electroencephalography (EEG) data. Furthermore, we test the algorithm on synthetic data and show the results of inverting real data.

2. Notation. We write $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ for the set of natural numbers including 0. Similarly, $\mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$ denotes the set of non-negative reals. We use $B_R \subset \mathbb{R}^3$ for the ball with radius R and centre 0 and $S \subset \mathbb{R}^3$ for the unit sphere in the three-dimensional space $\mathbb{R}^3$, whereas $B_R^{\text{ext}} := \mathbb{R}^3 \setminus B_R$ is the complement of a ball. The standard basis of $\mathbb{R}^3$ is denoted by ε^j, $j = 1, 2, 3$, that is, $\varepsilon^1 := (1, 0, 0)^\intercal$, etc.

The Euclidean inner product is denoted by $\cdot$, the vector product by $\times$, and the tensor product by $\otimes$. In addition, throughout the paper we will use the abbreviations $x = r\xi$, $y = s\eta$, and $z = v\zeta$ with unit vectors $\xi, \eta, \zeta \in S$ and radii $r = |x|$, $s = |y|$, and $v = |z|$. This decomposition is unique for all elements in $\mathbb{R}^3 \setminus \{0\}$.

In our notation, ∇_ξ^* denotes the part of the gradient ∇ containing the tangential derivatives multiplied by r; see [21, eq. (2.136)]. The operator ∇_ξ^* is often called the surface gradient. The differential operator L^* (independent of the radius r) is defined by

$$L_\xi^* := x \times \nabla_x = \xi \times \nabla_\xi^*$$

and called the surface curl operator. The Beltrami operator is given by $\Delta_\xi^* = \nabla_\xi^* \cdot \nabla_\xi^*$ (see [21, eq. (2.140)]) and is the part of the Laplacian independent of the radius. Note that variables as indices of operators indicate the dependence to which the operator is applied, since this is not always unique otherwise; see, e.g., equation (3.1) below.

3. Construction of suitable orthonormal basis functions. The RFMP uses a dictionary for approximating the unknown neuronal current. We will populate the dictionary with both global and localized functions. Both of these function types are built from an orthonormal basis adapted to the particular problem, which is also used in the singular value decomposition (SVD) of the MEG and EEG operators; see [35, 38].

The construction of our basis functions is based on spherical harmonics. Recall that a function of $\text{Harm}_n(S)$ (i.e., the space of all homogeneous, harmonic polynomials of degree n

ETNA
Kent State University and
Johann Radon Institute (RICAM)

restricted to the unit sphere S; see [21, Def. 3.22]) is called a spherical harmonic of degree $n \in \mathbb{N}_0$. By $\{Y_{n,j}\}_{j=1,\dots,2n+1}$, we denote an $\mathrm{L}^2(S)$-orthonormal set in $\mathrm{Harm}_n(S)$ (see [21, Rem. 3.25]), where $\mathrm{L}^2(S) \coloneqq \mathrm{L}^2(S,\mathbb{R})$. For more details on scalar spherical harmonics and their properties, see, for instance, [21]. For an introduction to (vector-valued) Lebesgue spaces, see [3].

By means of the scalar spherical harmonics, we can define a complete $\mathrm{L}^2(S,\mathbb{R}^3)$-orthonormal system of vector-valued spherical harmonics (see [21, Thm. 5.56]), which goes back to Edmonds (see [12]), and is defined (see [21, eqs. (5.309)–(5.311)]) by

$$\tilde{y}_{n,j}^{(i)}(\xi) \coloneqq (\tilde{\mu}_n^{(i)})^{-1/2}\tilde{o}_{n,\xi}^{(i)}Y_{n,j}(\xi), \quad \xi \in S,$$

where

$$\tilde{\mu}_n^{(i)} \coloneqq \begin{cases} (n+1)(2n+1), & \text{for } i=1,\\ n(2n+1), & \text{for } i=2,\\ n(n+1), & \text{for } i=3, \end{cases} \quad \tilde{o}_{n,\xi}^{(i)} \coloneqq \begin{cases} (n+1)\xi - \nabla_\xi^*, & \text{for } i=1,\\ n\xi + \nabla_\xi^*, & \text{for } i=2,\\ L_\xi^*, & \text{for } i=3, \end{cases}$$

for all $i \in \{1,2,3\}$, $n \in \mathbb{N}_{0_i}$, and $j = 1,\dots,2n+1$. We use $\mathbb{N}_{0_i}$ as an abbreviation for $\mathbb{N}_0$ in the case of $i=1$ and for $\mathbb{N}$ in the case of $i \in \{2,3\}$. Note that these vector-valued spherical harmonics are homogeneous harmonic polynomials. These and more information on vector spherical harmonics can be found in [21]. In analogy, we can define, for all $i \in \{1,2,3\}$ and $n \in \mathbb{N}_{0_i}$, Edmonds vector Legendre polynomials by means of the scalar Legendre polynomials P_n (see [21, eq. (3.165)]) of degree $n \in \mathbb{N}_0$ and type $i \in \{1,2,3\}$ as in [21, Lem. 5.63]:

$$\tilde{p}_n^{(i)}(\xi,\eta) \coloneqq (\tilde{\mu}_n^{(i)})^{-1/2}\tilde{o}_{n,\xi}^{(i)}P_n(\xi\cdot\eta), \quad \xi,\,\eta \in S. \tag{3.1}$$

Besides this, the Legendre polynomials and their vectorial counterparts possess addition theorems; see [21, Thms. 3.26 and 5.64]. These imply, for all types, degrees, and $\xi,\,\eta \in S$, the representations

$$\sum_{j=1}^{2n+1} Y_{n,j}(\xi)Y_{n,j}(\eta) = \frac{2n+1}{4\pi}P_n(\xi\cdot\eta), \quad \sum_{j=1}^{2n+1} \tilde{y}_{n,j}^{(i)}(\xi)Y_{n,j}(\eta) = \frac{2n+1}{4\pi}\tilde{p}_n^{(i)}(\xi,\eta). \tag{3.2}$$

Furthermore, we also obtain for (pointwise) Euclidean norms,

$$\sum_{j=1}^{2n+1} \left|\tilde{y}_{n,j}^{(i)}(\xi)\right|^2 = \frac{2n+1}{4\pi}, \quad \xi \in S.$$

Based on a separation ansatz, we combine the vector spherical harmonics with orthogonal Jacobi polynomials $P_m^{(\alpha,\beta)}$ for the radial part in order to achieve an orthonormal set of functions over the ball. For further details on Jacobi polynomials, see [56].

THEOREM 3.1. *The set of vector-valued functions* $\tilde{g}_{m,n,j}^{(i)}(R;\cdot)$ *for* $i \in \{1,2,3\}$, $m \in \mathbb{N}_0$, $n \in \mathbb{N}_{0_i}$, *and* $j = 1,\dots,2n+1$ *with the parameter*

$$t_n^{(i)} \coloneqq \begin{cases} n, & \text{for } i=1,\,3,\\ n-1, & \text{for } i=2, \end{cases}$$

is defined via

$$\tilde{g}_{m,n,j}^{(i)}(R;x) \coloneqq \sqrt{\frac{4m+2t_n^{(i)}+3}{R^3}}\left(\frac{r}{R}\right)^{t_n^{(i)}} P_m^{(0,t_n^{(i)}+1/2)}\left(2\frac{r^2}{R^2}-1\right)\tilde{y}_{n,j}^{(i)}(\xi), \quad x \in B_R.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

It is a complete orthonormal system in $\mathrm{L}^2(B_R, \mathbb{R}^3)$. *Using the abbreviation* $f^\wedge(i,m,n,j) \coloneqq \langle f, \tilde{g}^{(i)}_{m,n,j}(R;\cdot)\rangle_{\mathrm{L}^2(B_R,\mathbb{R}^3)}$, *each* $f \in \mathrm{L}^2(B_R,\mathbb{R}^3)$ *has the representation*

$$f = \sum_{i=1}^{3} \sum_{m=0}^{\infty} \sum_{n=0_i}^{\infty} \sum_{j=1}^{2n+1} f^\wedge(i,m,n,j)\tilde{g}^{(i)}_{m,n,j}(R;\cdot),$$

which converges unconditionally and strongly in the $\mathrm{L}^2(B_R,\mathbb{R}^3)$*-sense.*

For the derivation of this basis, see [35, 38, 46]. Recall that we showed in these papers that other choices for the parameter sequence $(t_n^{(i)})_n$ are also possible as long as the condition $\inf_{n\in\mathbb{N}_{0_i}} t_n^{(i)} \geq -3/2$ is satisfied. However, the sequence $(t_n^{(i)})_n$ stated in Theorem 3.1 is the natural choice for our application, as shown in [35]. For $x \in B_R$, a scalar-valued analogue can be defined via

$$G^{(i)}_{m,n,j}(R;x) \coloneqq \sqrt{\frac{4m+2t_n^{(i)}+3}{R^3}} \left(\frac{r}{R}\right)^{t_n^{(i)}} P_m^{(0,t_n^{(i)}+1/2)}\left(2\frac{r^2}{R^2}-1\right) Y_{n,j}(\xi). \tag{3.3}$$

4. Regularized functional matching pursuit. Before we apply the RFMP and its enhancement, the ROFMP, to the inverse MEG and EEG problem, we give a brief survey on the idea of the algorithm and the main theoretical results concerning the convergence of the algorithm. The problem can be described as follows: Let $\mathcal{X}$ and $\mathcal{Y}$ be two (possibly infinite-dimensional) Hilbert spaces. Let $g \in \mathcal{Y}$ be the given data (function) and $\mathcal{A}\colon \mathcal{X} \to \mathcal{Y}$ be a given linear and continuous operator. The task is to find a solution $f \in \mathcal{X}$ such that

$$\mathcal{A}f = g.$$

We proceed by considering the RFMP iteration in detail, which constructs a sequence $\{f_k\}_{k\in\mathbb{N}} \subset \mathcal{X}$ of approximations to the (unknown) solution f. The approximations f_k are linear combinations of elements from a so-called dictionary $\mathcal{D} \subset \mathcal{X}$ containing possibly useful trial functions. Thus, in the kth step we have already calculated the approximation f_k and we are searching for a tuple $(\alpha_{k+1}, d_{k+1}) \in \mathbb{R} \times \mathcal{D}$ containing the best-fitting next dictionary element and its coefficient. In this context, a best-fitting element is meant as a minimizer of the resulting regularized Tikhonov functional, that is,

$$\begin{aligned}(\alpha_{k+1}, d_{k+1}) &= \operatorname*{arg\,min}_{(\alpha,d)\in\mathbb{R}\times\mathcal{D}} \mathcal{J}_\lambda(g, f_k, d, \alpha),\\ \mathcal{J}_\lambda(g,f,d,\alpha) &\coloneqq \|g - \mathcal{A}(f+\alpha d)\|_{\mathcal{Y}}^2 + \lambda \|f+\alpha d\|_{\mathcal{X}}^2.\end{aligned} \tag{4.1}$$

Assume that the next chosen dictionary element d_{k+1} is already known. Then the optimal coefficient α_{k+1} can be calculated as the root of the derivative with respect to α of this regularized Tikhonov functional; see [15]. In addition, the second derivative with respect to α of the regularized Tikhonov functional is non-negative in general and positive for $d_{k+1} \in \mathcal{X} \setminus \ker \mathcal{A}$, that is,

$$\frac{\partial^2}{\partial\alpha^2}(\|g - \mathcal{A}(f_k + \alpha d_{k+1})\|_{\mathcal{Y}}^2 + \lambda \|f_k + \alpha d_{k+1}\|_{\mathcal{X}}^2) = 2\|\mathcal{A}d_{k+1}\|_{\mathcal{Y}}^2 + 2\lambda\|d_{k+1}\|_{\mathcal{X}}^2 \geq 0.$$

Thus, the regularized Tikhonov functional is convex as a function in α and the optimal coefficient α_{k+1} is unique. This greedy procedure leads to the next algorithm.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

ALGORITHM 4.1 (RFMP, [47, Algo. 2]). *Choose a dictionary $\mathcal{D} \subset \mathcal{X} \setminus \{0\}$, an initial approximation $f_0 \in \mathcal{X}$ (for example, $f_0 = 0$), and a regularization parameter $\lambda \in \mathbb{R}^+$.*

1. *Initialize the step counter to $k := 0$, define the residual $r_0 := g - \mathcal{A}f_0$, and choose a stopping criterion.*
2. *Find*

$$d_{k+1} := \operatorname*{arg\,max}_{d \in \mathcal{D}} \frac{(\langle r_k, \mathcal{A}d\rangle_{\mathcal{Y}} - \lambda\langle f_k, d\rangle_{\mathcal{X}})^2}{\|\mathcal{A}d\|_{\mathcal{Y}}^2 + \lambda\|d\|_{\mathcal{X}}^2} \tag{4.2}$$

   and set

$$\alpha_{k+1} := \frac{\langle r_k, \mathcal{A}d_{k+1}\rangle_{\mathcal{Y}} - \lambda\langle f_k, d_{k+1}\rangle_{\mathcal{X}}}{\|\mathcal{A}d_{k+1}\|_{\mathcal{Y}}^2 + \lambda\|d_{k+1}\|_{\mathcal{X}}^2},$$

   as well as $f_{k+1} := f_k + \alpha_{k+1}d_{k+1}$ and $r_{k+1} := g - \mathcal{A}f_{k+1} = r_k - \alpha_{k+1}\mathcal{A}d_{k+1}$.
3. *If the stopping criterion is fulfilled, then f_{k+1} is the output. Otherwise, increase k by 1 and go to step 2.*

Hence, in each step of the algorithm, the next dictionary element must be chosen and the corresponding coefficient must be calculated. In both cases, common expressions such as inner products between the dictionary elements or the forward operator action $\mathcal{A}d$ for all dictionary elements $d \in \mathcal{D}$ arise. These quantities can be calculated in a preprocessing step in order to accelerate the algorithm.

In practice, the algorithm will be stopped when some stopping criterion is satisfied; see [44]. The stopping criterion can, for instance, be a (relative) bound for the residual or an upper bound for the number of iterations. However, we neglect the stopping criterion for the analysis of the convergence of the sequence produced by the algorithm.

Several properties required for the convergence theorem of the RFMP are analysed in [15, 35, 44]. Note that in earlier publications the L^2-space is used instead of a general Hilbert space $\mathcal{X}$. The proofs are, however, easily transferable to the general case, as demonstrated in [47] for the convergence theorem.

The following theorem improves former results from [15, 16, 44, 49]. In [47], the result is formulated for the case of a finite-dimensional Hilbert space $\mathcal{Y}$. However, the proof of this result is in fact independent of the structure of the operator range and the space $\mathcal{Y}$. Thus, the next theorem is also valid in the context of infinite-dimensional Hilbert spaces $\mathcal{Y}$. The latest approach to the convergence results can be found in [33, 34, 45].

THEOREM 4.2 (Convergence theorem, [45]). *Let an arbitrary regularization parameter $\lambda \in \mathbb{R}^+$ be given. If the sequence $\{f_k\}_{k\in\mathbb{N}}$ is produced by the RFMP, then $\{f_k\}_{k\in\mathbb{N}}$ converges in $\mathcal{X}$ to $f_\infty := f_0 + \sum_{k=1}^{\infty} \alpha_k d_k \in \mathcal{X}$ and the sequence of residuals $\{r_k\}_{k\in\mathbb{N}}$ converges in $\mathcal{Y}$. Moreover, if the span of the dictionary is dense in $\mathcal{X}$ or $(\ker \mathcal{A})^\perp$, then for all $g^\delta \in \mathcal{Y}$, the function f_∞ obtained by the RFMP solves the Tikhonov-regularized normal equation*

$$(\mathcal{A}^*\mathcal{A} + \lambda\mathcal{I})f_\infty = \mathcal{A}^* g^\delta, \tag{4.3}$$

where $\mathcal{A}^*$ is the adjoint operator of $\mathcal{A}$ and $\mathcal{I}$ is the identity operator on $\mathcal{X}$.

Proof. The case $\overline{\operatorname{span}\mathcal{D}} = \mathcal{X}$, where the closure of the span of the dictionary equals the Hilbert space $\mathcal{X}$, is proven in [45, Thms. 5.4.10 and 5.4.11] and [47, Thm. 4], where the latter is the idea generator for this proof. Since $(\ker \mathcal{A})^\perp$ is a closed subspace of the Hilbert space $\mathcal{X}$, the space $((\ker \mathcal{A})^\perp, \langle\cdot,\cdot\rangle_{\mathcal{X}})$ is a (complete) Hilbert space. Hence, if $\overline{\operatorname{span}\mathcal{D}} = (\ker \mathcal{A})^\perp$, then we can replace $\mathcal{X}$ by $(\ker \mathcal{A})^\perp$ in the known propositions and $f_{\infty,(\ker \mathcal{A})^\perp}$ exists as the limit of $(f_k)_k$ produced by the RFMP in $(\ker \mathcal{A})^\perp$. We define the auxiliary operators

$$\mathcal{A}_{(\ker \mathcal{A})^\perp} := \mathcal{A}\mathcal{P}_{(\ker \mathcal{A})^\perp} : \mathcal{X} \to \mathcal{Y},$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$$\mathcal{A}^*_{(\ker \mathcal{A})^\perp} = (\mathcal{A}\mathcal{P}_{(\ker \mathcal{A})^\perp})^* = \mathcal{P}_{(\ker \mathcal{A})^\perp}\mathcal{A}^* = \mathcal{A}^*,$$

where the last equation is due to the fact that $\overline{\operatorname{ran} \mathcal{A}^*} = (\ker \mathcal{A})^\perp$. We get that $f_{\infty,(\ker \mathcal{A})^\perp}$ solves the Tikhonov-regularized normal equation in $((\ker \mathcal{A})^\perp, \langle\cdot,\cdot\rangle_\mathcal{X})$, that is,

$$(\mathcal{A}^*\mathcal{A}_{(\ker \mathcal{A})^\perp} + \lambda\mathcal{I}_{(\ker \mathcal{A})^\perp})f_{\infty,(\ker \mathcal{A})^\perp} = \mathcal{A}^* g^\delta.$$

In addition, we have $\mathcal{A}f = \mathcal{A}_{(\ker \mathcal{A})^\perp} f$ for all $f \in \mathcal{X}$. Thus, we arrive at

$$(\mathcal{A}^*\mathcal{A} + \lambda\mathcal{I}_{(\ker \mathcal{A})^\perp})f_{\infty,(\ker \mathcal{A})^\perp} = \mathcal{A}^* g^\delta.$$

Since $\mathcal{I}_{(\ker \mathcal{A})^\perp} f_{\infty,(\ker \mathcal{A})^\perp} = f_{\infty,(\ker \mathcal{A})^\perp}$, we arrive at the desired result. □

It is a basic result of Tikhonov regularization that every solution of equation (4.3) minimizes the regularized Tikhonov–Philips functional. Since $\lambda > 0$, this minimizer and the solution of the Tikhonov-regularized normal equation are both uniquely determined by

$$f_\infty = (\mathcal{A}^*\mathcal{A} + \lambda\mathcal{I})^{-1}\mathcal{A}^* g^\delta, \quad g^\delta \in \mathcal{Y}.$$

If the smoothness of the function f is unknown, it may be necessary to regularize with certain different norms that are generated by a pseudo-differential operator; see [39, Ch. 4.2]. By means of the pseudo-differential operators, we are able to define a function space similar to the Sobolev spaces on the ball.

DEFINITION 4.3. *Let $\mathcal{B}\colon (\ker \mathcal{A})^\perp \to \mathcal{X}$ be a densely defined pseudo-differential operator with $\|\mathcal{B}f\|_\mathcal{X} \geq \beta\|f\|_\mathcal{X}$ for a $\beta > 0$ and all $f \in \operatorname{dom}\mathcal{B}$. Furthermore, let its singular system $(x_k, y_k; \beta_k)$ be given such that*

$$\mathcal{B}f = \sum_{k=1}^{\infty} \beta_k \langle f, x_k\rangle_\mathcal{X} y_k.$$

With the definition

$$\mathcal{E} := \left\{ f \in (\ker \mathcal{A})^\perp \;\middle|\; \sum_{k=1}^{\infty} \beta_k^2 \langle f, x_k\rangle_\mathcal{X}^2 < \infty \right\},$$

we get $\operatorname{dom}\mathcal{B} = \mathcal{E}$, which is dense in $\mathcal{X}$ with respect to the $\mathcal{X}$-norm. In addition, based on the inner product

$$\langle f, h\rangle_\mathcal{H} := \sum_{k=1}^{\infty} \beta_k^2 \langle f, x_k\rangle_\mathcal{X}\langle h, x_k\rangle_\mathcal{X} = \langle \mathcal{B}f, \mathcal{B}h\rangle_\mathcal{X},$$

we define $\mathcal{H}$ as the completion of $\mathcal{E}$.

Next, we use the Hilbert space generated by the pseudo-differential operator as the penalty term in the Tikhonov functional.

THEOREM 4.4. *Let $\mathcal{H} \subset \mathcal{X}$ as in Definition 4.3, and let $g^\delta \in \mathcal{Y}$. Then, the solution of the RFMP with respect to the Tikhonov functional*

$$\mathcal{J}_\lambda(f) = \left\|\mathcal{A}f - g^\delta\right\|_\mathcal{Y}^2 + \lambda\left\|f\right\|_\mathcal{H}^2$$

converges to $f_\infty \in \mathcal{H}$ fulfilling

$$(\mathcal{P}_\mathcal{H}\mathcal{A}^*\mathcal{A} + \lambda\mathcal{B}^*\mathcal{B})f_\infty = \mathcal{P}_\mathcal{H}\mathcal{A}^* g^\delta.$$

Note that $\mathcal{B}^*$ is the adjoint operator of $\mathcal{B}$ with respect to the topology in $\mathcal{X}$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proof. In order to apply Theorem 4.2 to the problem $(\mathcal{A}_{\mathcal{H}}, \mathcal{H}, \mathcal{Y})$ with $\mathcal{A}_{\mathcal{H}} := \mathcal{A}|_{\mathcal{H}}$, we need the continuity of $\mathcal{A}_{\mathcal{H}}$, which is given by

$$\begin{aligned}\|\mathcal{A}_{\mathcal{H}}\|_{\mathcal{L}(\mathcal{H},\mathcal{Y})} &= \sup_{f\in\mathcal{H},f\neq 0} \frac{\|\mathcal{A}_{\mathcal{H}} f\|_{\mathcal{Y}}}{\|f\|_{\mathcal{H}}} = \sup_{f\in\mathcal{H},f\neq 0} \frac{\|\mathcal{A} f\|_{\mathcal{Y}}}{\|f\|_{\mathcal{H}}} \\ &\leq \sup_{f\in\mathcal{H},f\neq 0} \frac{\|\mathcal{A}\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})}\|f\|_{\mathcal{X}}}{\|f\|_{\mathcal{H}}} \leq \frac{1}{\beta}\|\mathcal{A}\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})} < \infty.\end{aligned}$$

Here, we have used that $\mathcal{A}\colon \mathcal{X} \to \mathcal{Y}$ is a bounded linear operator and $\|f\|_{\mathcal{H}} \geq \beta\|f\|_{\mathcal{X}}$ for a $\beta > 0$. Thus, the sequence produced by the RFMP converges to $f_\infty \in \mathcal{H}$, which fulfils

$$(\mathcal{A}_{\mathcal{H}}^*\mathcal{A}_{\mathcal{H}} + \lambda\mathcal{I}_{\mathcal{H}})f_\infty = \mathcal{A}_{\mathcal{H}}^* g^\delta; \tag{4.4}$$

see Theorem 4.2. Note that the adjoint operator $\mathcal{A}_{\mathcal{H}}^*$ of $\mathcal{A}_{\mathcal{H}}$ is subject to the $\mathcal{H}$-inner product. For all $f \in \mathcal{H}$, $g^\delta \in \mathcal{Y}$, we also get the relations

$$\begin{aligned}\big\langle f, \mathcal{A}_{\mathcal{H}}^* g^\delta\big\rangle_{\mathcal{H}} &= \big\langle \mathcal{A}_{\mathcal{H}} f, g^\delta\big\rangle_{\mathcal{Y}} = \big\langle \mathcal{A} f, g^\delta\big\rangle_{\mathcal{Y}} = \big\langle f, \mathcal{A}^* g^\delta\big\rangle_{\mathcal{X}}, \\ \big\langle f, \mathcal{A}_{\mathcal{H}}^* g^\delta\big\rangle_{\mathcal{H}} &= \big\langle \mathcal{B} f, \mathcal{B}\mathcal{A}_{\mathcal{H}}^* g^\delta\big\rangle_{\mathcal{X}} = \big\langle f, \mathcal{B}^*\mathcal{B}\mathcal{A}_{\mathcal{H}}^* g^\delta\big\rangle_{\mathcal{X}}.\end{aligned}$$

Thus,

$$\big\langle f, \mathcal{B}^*\mathcal{B}\mathcal{A}_{\mathcal{H}}^* g^\delta - \mathcal{A}^* g^\delta\big\rangle_{\mathcal{X}} = 0 \quad \text{for all } f \in \mathcal{H},$$

which implies $(\mathcal{B}^*\mathcal{B}\mathcal{A}_{\mathcal{H}}^* - \mathcal{P}_{\mathcal{H}}\mathcal{A}^*)g^\delta \in \mathcal{H}^\perp = \{0\}$ since $\mathcal{H}$ is dense in $\mathcal{X}$ and $\operatorname{ran}\mathcal{B}^* = (\ker\mathcal{B})^\perp = \operatorname{dom}\mathcal{B} \subset \mathcal{H}$. Accordingly, we get $\mathcal{B}^*\mathcal{B}\mathcal{A}_{\mathcal{H}}^* = \mathcal{P}_{\mathcal{H}}\mathcal{A}^*$ because g^δ can be arbitrary. Due to its definition, the operator $\mathcal{B}^*\mathcal{B}$ has a bounded inverse. Thus, for the adjoint operator $\mathcal{A}_{\mathcal{H}}^*$, the identity $\mathcal{A}_{\mathcal{H}}^* = (\mathcal{B}^*\mathcal{B})^{-1}\mathcal{P}_{\mathcal{H}}\mathcal{A}^*$ holds true. Inserting this into equation (4.4), we immediately get

$$\begin{aligned}&((\mathcal{B}^*\mathcal{B})^{-1}\mathcal{P}_{\mathcal{H}}\mathcal{A}^*\mathcal{A}_{\mathcal{H}} + \lambda\mathcal{I}_{\mathcal{H}})f_\infty = (\mathcal{B}^*\mathcal{B})^{-1}\mathcal{P}_{\mathcal{H}}\mathcal{A}^* g^\delta, \\ \iff\quad &(\mathcal{P}_{\mathcal{H}}\mathcal{A}^*\mathcal{A} + \lambda\mathcal{B}^*\mathcal{B})f_\infty = \mathcal{P}_{\mathcal{H}}\mathcal{A}^* g^\delta. \qquad \square\end{aligned}$$

If more knowledge on the operator $\mathcal{A}$ is available, for instance an SVD $\{x_k, y_k; \sigma_k\}_{k\in\mathbb{N}}$, then the spanning condition for the dictionary in Theorem 4.2 can be weakened.

THEOREM 4.5 ([47, Thm. 6]). *Let additionally $\mathcal{A}$ be assumed to be a compact operator with singular system $\{x_k, y_k; \sigma_k\}_{k\in\mathbb{N}}$. Let the conditions of Theorem 4.2 be fulfilled, except that the dictionary is only a spanning set for $\mathcal{V} := \overline{\operatorname{span}\{x_k\}_{k\in\mathcal{J}}}$, where $\mathcal{J} \subset \mathbb{N}$ is a countable index set.*

Then the solution $f_{\infty,\mathcal{V}} \in \mathcal{V}$ produced by the RFMP and the unique solution of the Tikhonov-regularized normal equation $f_\infty \in \mathcal{X}$ satisfy

$$(\mathcal{A}^*\mathcal{A} + \lambda\mathcal{I})f_{\infty,\mathcal{V}} = (\mathcal{A}^*\mathcal{A} + \lambda\mathcal{I})\mathcal{P}_{\mathcal{V}} f_\infty.$$

Since $\lambda > 0$, the operator $\mathcal{A}^*\mathcal{A} + \lambda\mathcal{I}$ is one-to-one and we obtain $f_{\infty,\mathcal{V}} = \mathcal{P}_{\mathcal{V}} f_\infty$.

For a fixed regularization parameter $\lambda > 0$, we have already seen that the limit of the sequence $\{f_k\}_{k\in\mathbb{N}}$ of approximations obtained by the RFMP converges to f_∞, which is the unique minimizer of the Tikhonov–Philips functional. In [33], this is generalized to our setting under appropriate assumptions. Therein, it is additionally taken into account that the RFMP stops after a finite number of iterations K.

THEOREM 4.6 ([33, Thm. 9.35]). *Let $\{f_{\lambda,k}^\delta\}_{k\in\mathbb{N}_0}$ be the sequence of iterations of the RFMP to the inverse problem $\mathcal{A}f = g^\delta$ with $g^\delta \in \mathcal{Y}$ fulfilling $\|g - g^\delta\|_{\mathcal{Y}} \leq \delta$ using the regularization parameter $\lambda > 0$. We assume that*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

- *the dictionary is normalized in the sense that* $\|\mathcal{A}d\|_{\mathcal{Y}}^2 + \|d\|_{\mathcal{X}}^2 = 1$ *for all* $d \in \mathcal{D}$,
- *the best approximate solution* f^+ *fulfils a Hölder-type source condition for* $\nu = 1$,
- *the function* $|f_{\lambda,k}^\delta|_{\mathcal{D}}$ *is uniformly bounded with respect to* λ, *where* $|\cdot|_{\mathcal{D}}$ *is a certain measure for the sparsity (see [33, Lem. 9.30]) and*
- *there exist constants* m_1, $m_2 > 0$ *such that* $\lambda(\delta) = m_1\delta^{2/3}$ *and* $K(\delta) = m_2\delta^{-6}$.

Then there exists a constant $C > 0$ *such that*

$$\left\| f_{\lambda,K(\delta)}^\delta - f^+ \right\|_{\mathcal{X}} \leq C\delta^{2/3}.$$

Eventually, this theorem shows that the RFMP has a convergence rate of $\delta^{2/3}$ for $\delta \to 0+$ and that this rate can even be retained when only a finite number of iterations is used; see [33, p. 160]. This convergence rate is not surprising, since the order of convergence of the Tikhonov regularization is also $\delta^{2/3}$, which is optimal with respect to the source condition.

A problem of the RFMP occurring in numerical tests is that certain dictionary elements may be chosen repeatedly. This can be understood as a correction for some previously chosen coefficients $\alpha_1, \ldots, \alpha_k$. A reason for this phenomenon is that the residual r_k is not orthogonal to the image of the span of the previously chosen dictionary elements, that is, $\operatorname{span}\{\mathcal{A}d_1, \ldots, \mathcal{A}d_k\}$. In order to get rid of this unwanted effect, an enhancement of the RFMP is developed in [50, 57]. Therein, an orthogonalization step based on the idea of [52, 58] is introduced. The final algorithm is called the regularized orthogonal functional matching pursuit algorithm (ROFMP).

ALGORITHM 4.7 (ROFMP, [20, Algo. 8]). *Let* $\mathcal{A}$ *and* g *be given as in equation (4.1) with* $\mathcal{Y} := \mathbb{R}^\ell$ *and* $\ell \in \mathbb{N}$. *Choose a stopping criterion, a dictionary* $\mathcal{D} \subset \mathcal{X}$, *an initial approximation* $f_0 \in \mathcal{X}$, *and a regularization parameter* $\lambda \in \mathbb{R}^+$.

1. *Initialize the step counter to* $k := 0$, *define the residual* $r_0 := g - \mathcal{A}f_0$, *set* $\mathcal{V}_0 := \emptyset$, $\mathcal{W}_0 := \mathcal{V}_0^\perp$, *and* $\mathcal{B}_0(d) := \mathcal{I}_{\mathcal{X}}$.
2. *Find*

$$d_{k+1} := \operatorname*{arg\,max}_{d\in\mathcal{D}} \frac{\left(\langle r_k, \mathcal{P}_{\mathcal{W}_k}(\mathcal{A}d)\rangle_{\mathbb{R}^\ell} + \lambda\langle f_k, \mathcal{B}_k(d) - d\rangle_{\mathcal{X}}\right)^2}{\|\mathcal{P}_{\mathcal{W}_k}(\mathcal{A}d)\|_{\mathbb{R}^\ell}^2 + \lambda\|\mathcal{B}_k(d) - d\|_{\mathcal{X}}^2}$$

and set

$$\alpha_{k+1} := \frac{\langle r_k, \mathcal{P}_{\mathcal{W}_k}(\mathcal{A}d_{k+1})\rangle_{\mathbb{R}^\ell} + \lambda\langle f_k, \mathcal{B}_k(d_{k+1}) - d_{k+1}\rangle_{\mathcal{X}}}{\|\mathcal{P}_{\mathcal{W}_k}(\mathcal{A}d_{k+1})\|_{\mathbb{R}^\ell}^2 + \lambda\|\mathcal{B}_{k+1}(d_{k+1}) - d_{k+1}\|_{\mathcal{X}}^2}.$$

3. *For all* $d \in \mathcal{D}$, *define the mappings*

$$\begin{aligned}
\beta_k^{(k)}(d) &:= \frac{\langle \mathcal{A}d, \mathcal{P}_{\mathcal{W}_{k-1}}\mathcal{A}d_k\rangle_{\mathbb{R}^\ell}}{\|\mathcal{P}_{\mathcal{W}_{k-1}}\mathcal{A}d_k\|_{\mathbb{R}^\ell}^2}, \\
\beta_i^{(k)}(d) &= \beta_i^{(k-1)}(d) - \beta_k^{(k)}(d)\beta_i^{(k-1)}(d_k), \quad \text{for } i = 1, \ldots, k-1, \\
\mathcal{B}_k(d) &:= \sum_{i=1}^{k} \beta_i^{(k)}(d) d_i.
\end{aligned}$$

4. *Update the coefficients as follows:*

$$\begin{aligned}
\alpha_i^{(k+1)} &= \alpha_i^{(k)} - \alpha_{k+1}\beta_i^{(k)}(d_{k+1}), \quad \text{for } i = 1, \ldots, k, \\
\alpha_{k+1}^{(k+1)} &= \alpha_{k+1},
\end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

and set $f_{k+1} \coloneqq \sum_{i=1}^{k+1} \alpha_i^{(k+1)} d_i$ *as well as* $r_{k+1} \coloneqq g - \mathcal{A} f_{k+1}$. *Update the spaces such that*

$$\mathcal{V}_{k+1} \coloneqq \operatorname{span}\{\mathcal{A}d_1, \dots, \mathcal{A}d_{k+1}\}, \quad \mathcal{W}_{k+1} = \mathcal{V}_{k+1}^{\perp}.$$

5. *If the stopping criterion is satisfied, then use* f_{k+1} *as an approximate solution to* $\mathcal{A}f = g$. *Otherwise, increase* k *by* 1 *and go to step 2.*

Note that the recurrence formula for the $\beta_i^{(k)}$ can be resolved to get a closed formula; see [54, Thm. 4.4.3].

In order to obtain convergence of the ROFMP, an iterative refinement is necessary. In this case, after a certain number of steps $K \in \mathbb{N}$, the algorithm is restarted with $g - \mathcal{A}f_K$ as the new data vector. This allows us, in particular, to use the entire dictionary again after K iterations. It may be useful to keep the previous approximation in the penalty term; see [50].

5. The inverse electro-magnetoencephalography problem. For the inverse electro-magnetoencephalography problem, one is interested in the reconstruction of the human neuronal current from indirect measurements outside the head, namely the electric potential differences on the scalp and the magnetic flux density at the MEG sensors. The flux density is the part of the magnetic field pointing towards the normal direction ν of the sensor surface. However, this problem cannot be analysed in detail without modelling the head appropriately. Within our approach, we use the common *multiple-shell model*; see [10, 38]. For our numerical tests, we use the particular case of the three-shell model even though the model can accommodate an arbitrary number of shells $L \in \mathbb{N}$. In detail, we assume that

- the cerebrum is a closed ball with radius ϱ_0, that is, B_{ϱ_0},
- surrounding the cerebrum, there are $L \geq 2$ spherical shells $S_{[\varrho_l, \varrho_{l+1}]} \coloneqq \overline{B_{\varrho_{l+1}} \setminus B_{\varrho_l}}$, for $l = 0, \dots, L-1$, modelling the various head tissues,
- each tissue (i.e., each shell $S_{[\varrho_l, \varrho_{l+1}]}$ and the cerebrum B_{ϱ_0}) has a constant conductivity $\sigma_l > 0$ for all $l = 0, \dots, L-1$, and outside the head the conductivity is vanishing, i.e., $\sigma_L = 0$,
- the permeability is constant everywhere and equals the permeability of vacuum, μ_0,
- the relation between the neuronal current and the induced quantities can be modelled by means of quasi-static Maxwell's equations (see [53]), and
- the continuously distributed neuronal current $J \in \mathrm{L}^2(B_{\varrho_0}, \mathbb{R}^3)$ is non-vanishing only inside the cerebrum.

Note that we do not assume further smoothness or boundary conditions for the neuronal current.

The relation between the sought current and the measurements is, hence, given via partial differential equations. Starting from there, a derivation of Fredholm integral equations of the first kind is detailed in [38]. We, eventually, obtain the functionals

$$\begin{aligned} \mathcal{A}_{\mathrm{M}}^{\ell} J &= \nu(y_\ell) \cdot \left(\mu_0 \nabla_y \int_{B_{\varrho_0}} J(x) \cdot k_{\mathrm{M}}(x, y) \, \mathrm{d}x \right) \Big|_{y = y_\ell}, && \ell = 1, \dots, \ell_{\mathrm{M}}, \\ \mathcal{A}_{\mathrm{E}}^{\ell} J &= \int_{B_{\varrho_0}} J(x) \cdot k_{\mathrm{E}}(x, y_{\ell_{\mathrm{M}}+k}) \, \mathrm{d}x, && \ell = \ell_{\mathrm{M}} + 1, \dots, \ell_{\mathrm{M}} + \ell_{\mathrm{E}}, \end{aligned} \tag{5.1}$$

and associated operators

$$\left(\mathcal{A}_{\mathrm{M}} J\right)(y) = \mu_0 \nabla_y \int_{B_{\varrho_0}} J(x) \cdot k_{\mathrm{M}}(x, y) \, \mathrm{d}x, \quad y \in B_{\varrho_L}^{\mathrm{ext}}, \tag{5.2}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$$(\mathcal{A}_{\mathrm{E}} J)(y) = \int_{B_{\varrho_0}} J(x) \cdot k_{\mathrm{E}}(x,y)\,\mathrm{d}x, \qquad y \in B_{\varrho_L}^{\mathrm{ext}}. \tag{5.3}$$

We refer to [38] for precise representations of the occurring integral kernels $k_{\mathrm{M}} \in \mathrm{L}^2(B_{\varrho_0} \times B_{\varrho_L}^{\mathrm{ext}}, \mathbb{R}^3)$ and $k_{\mathrm{E}} \in \mathrm{L}^2(B_{\varrho_0} \times B_{\varrho_L}^{\mathrm{ext}}, \mathbb{R}^3)$. In the MEG case, the above considered functionals $\mathcal{A}_{\mathrm{M}}^{\ell}\colon \mathrm{L}^2(B_{\varrho_0}, \mathbb{R}^3) \to \mathbb{R}$ map the neuronal current onto the magnetic flux density evaluated at the sensor positions $y_\ell \in \overline{\mathbb{R}^3 \setminus B_{\varrho_L}}$, $\ell = 1, \ldots, \ell_{\mathrm{M}}$, outside the head. In the EEG case, the functionals $\mathcal{A}_{\mathrm{E}}^{\ell}\colon \mathrm{L}^2(B_{\varrho_0}, \mathbb{R}^3) \to \mathbb{R}$, $\ell = \ell_{\mathrm{M}} + 1, \ldots, \ell_{\mathrm{M}} + \ell_{\mathrm{E}}$, map the current onto the electric potential difference measured at several positions on the scalp.

The reconstruction of the neuronal currents is, in both cases, a severely ill-posed inverse problem; see [38]. Besides singular values decreasing exponentially fast towards zero, both problems deal with an infinite-dimensional operator null space. It was proven in [38] that only the harmonic and solenoid part of the neuronal current affects the measurement.

THEOREM 5.1 (Operator null space [38, Thms. 6.1 and 6.2]). *Let $J \in \mathrm{L}^2(B_{\varrho_0}, \mathbb{R}^3)$ be the neuronal current and let the MEG forward operator $\mathcal{A}_{\mathrm{M}}\colon \mathrm{L}^2(B_{\varrho_0}, \mathbb{R}^3) \to \mathrm{L}^2(B_{\varrho_L}^{\mathrm{ext}})$ be given by equation (5.2). On the one hand, the null space of $\mathcal{A}_{\mathrm{M}}$ is characterized by*

$$(\ker \mathcal{A}_{\mathrm{M}})^{\perp} = \overline{\operatorname{span}\{\tilde{g}_{0,n,j}^{(3)}(\varrho_0;\cdot) \mid n \in \mathbb{N}, j = 1, \ldots, 2n+1\}}.$$

Hence, only the harmonic part of the toroidal direction of the neuronal current is not in the operator null space. On the other hand, the null space of $\mathcal{A}_{\mathrm{E}}\colon \mathrm{L}^2(\mathbb{B}_{\varrho_0}, \mathbb{R}^3) \to \mathrm{L}^2(\mathbb{S}_{[\varrho_{L-1}, \varrho_L]})$, which is defined in equation (5.3), is characterized by

$$(\ker \mathcal{A}_{\mathrm{E}})^{\perp} = \overline{\operatorname{span}\{\tilde{g}_{0,n,j}^{(2)}(\varrho_0;\cdot) \mid n \in \mathbb{N}, j = 1, \ldots, 2n+1\}}.$$

Thus, only the harmonic part of directions related to inner harmonics of the neuronal current is not in the operator null space.

In addition, the part of the neuronal current that can be measured via the MEG device is silent to the EEG measurement and vice versa. Taking this into account, we are able to gain more information about the neuronal current by measuring the electric potential and magnetic flux induced by the brain activity simultaneously. To be more precise, the spaces $(\ker \mathcal{A}_{\mathrm{M}})^{\perp}$ and $(\ker \mathcal{A}_{\mathrm{E}})^{\perp}$ are orthogonal with respect to $\mathrm{L}^2(B_{\varrho_0}, \mathbb{R}^3)$. Note that joint inversions, where usually a part of the solution is jointly influenced by different types of data, are connected to typical problems (such as weighting heterogeneous data; see, e.g., [32]). However, in our particular case, we need not deal with such problems due to the complementarity of these spaces. For this reason, we will perform separate inversions of the MEG and the EEG data, which, as an additional benefit, also reduces the dimension of the inverse problem(s) to be solved. In the next section, we show that this can also theoretically be supported with respect to the used algorithm.

6. Simultaneous but complementary inversion. The RFMP can be used for a joint inversion of several data types in order to yield more information about the source. This may be reasonable if a source induces several physical quantities that can be measured simultaneously, as in the case of the reconstruction of neuronal currents or the joint gravitation and normal-mode inversion; see [15] for the latter. However, due to the structure of the MEG and EEG null spaces, we are able to prove that a joint inversion cannot yield more details of the neuronal current than independent single inversions.

For this purpose, we assume that the Hilbert space of the operator domain $\mathcal{X}$ can be divided into two orthogonal subspaces $\mathcal{X}_{\mathrm{M}}$ and $\mathcal{X}_{\mathrm{E}}$ such that $\mathcal{X} = \mathcal{X}_{\mathrm{M}} \oplus \mathcal{X}_{\mathrm{E}}$. Now, we introduce the abbreviation $\bullet \in \{\mathrm{M}, \mathrm{E}\}$ that will be used if a quantity can be used for either the MEG or the EEG case.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

DEFINITION 6.1 (Joint inversion). *In all three cases, we want to find the minimizer of equation (4.1) by means of the RFMP over the respective Hilbert spaces. In addition, we choose the regularization parameter* $\lambda := (\lambda_{\mathrm{M}}, \lambda_{\mathrm{E}})^{\mathsf{T}}$ *for the joint inversion with regularization parameters* $\lambda_{\mathrm{M}}, \lambda_{\mathrm{E}} > 0$ *for the MEG and EEG inversion.*

Separate: Let $\mathcal{X}_\bullet$, $\mathcal{Y}_\bullet := \mathbb{R}^{\ell_\bullet}$ *with an* $\ell_\bullet \in \mathbb{N}$, *and the data vector* $g_\bullet \in \mathbb{R}^{\ell_\bullet}$ *be given. Then, the forward operator* $\mathcal{A}_\bullet \colon \mathcal{X}_\bullet \to \mathbb{R}^{\ell_\bullet}$ *is considered. In addition, let the span of the dictionary* $\mathcal{D}_\bullet$ *be dense in* $(\ker \mathcal{A}_\bullet)^\perp$.

Joint: Let $\mathcal{X}$ *be defined by the internal direct sum, that is,* $\mathcal{X} := \mathcal{X}_{\mathrm{M}} \oplus \mathcal{X}_{\mathrm{E}}$, *let* $\mathcal{Y} := \mathbb{R}^{\ell_\Sigma}$ *with* $\ell_\Sigma := \ell_{\mathrm{M}} + \ell_{\mathrm{E}}$, *let the data be given by* $g^{\mathsf{T}} := (g_{\mathrm{M}}^{\mathsf{T}}, g_{\mathrm{E}}^{\mathsf{T}}) \in \mathbb{R}^{\ell_\Sigma}$, *and let the operator* $\mathcal{A} \colon \mathcal{X} \to \mathbb{R}^{\ell_\Sigma}$ *for the RFMP be given by*

$$\mathcal{A}J := \left(\left(\mathcal{A}_{\mathrm{M}} \left(\mathcal{P}_{\mathcal{X}_{\mathrm{M}}} J \right) \right)^{\mathsf{T}}, \left(\mathcal{A}_{\mathrm{E}} \left(\mathcal{P}_{\mathcal{X}_{\mathrm{E}}} J \right) \right)^{\mathsf{T}} \right)^{\mathsf{T}}, \tag{6.1}$$

where $\mathcal{P}_{\mathcal{X}_\bullet}$ *are the orthogonal projections onto the respective spaces. For the dictionary, we choose* $\mathcal{D} = \mathcal{D}_{\mathrm{M}} \cup \mathcal{D}_{\mathrm{E}}$. *Moreover, we replace the Tikhonov functional (4.1) by*

$$\|g - \mathcal{A}(f + \alpha d)\|^2_{\mathbb{R}^{\ell_\Sigma}} + \lambda_{\mathrm{M}} \|\mathcal{P}_{\mathcal{X}_{\mathrm{M}}}(f + \alpha d)\|^2_{\mathcal{X}_{\mathrm{M}}} + \lambda_{\mathrm{E}} \|\mathcal{P}_{\mathcal{X}_{\mathrm{E}}}(f + \alpha d)\|^2_{\mathcal{X}_{\mathrm{E}}}. \tag{6.2}$$

In this setting, we can prove that no more information can be obtained by a simultaneous joint inversion based on the RFMP than by summing the two solutions of independent inversion of both data sets. This does not contradict former results achieved by the methodology of multiparameter regularization (see, e.g., [7] and references therein), since this result is directly related to the orthogonality structure of the considered spaces.

THEOREM 6.2. *Let the setting of Definition 6.1 be given, and let fixed regularization parameters* $\lambda_{\mathrm{M}} > 0$ *and* $\lambda_{\mathrm{E}} > 0$ *be chosen. Then, the RFMP solution of the joint case* f_∞ *coincides with the sum of the solutions independently obtained by the RFMP in the separate cases, that is,* $f_\infty = f_{\mathrm{M},\infty} + f_{\mathrm{E},\infty}$.

Proof. Note that Definition 6.1 implies all the conditions of Theorem 4.2. Thus, we immediately obtain that the two solutions of the independent cases fulfil

$$(\mathcal{A}_\bullet^* \mathcal{A}_\bullet + \lambda_\bullet \mathcal{I}_{\mathcal{X}_\bullet}) f_{\bullet,\infty} = \mathcal{A}_\bullet^* g_\bullet, \quad \text{in } \mathcal{X}_\bullet. \tag{6.3}$$

In addition, $\overline{\operatorname{ran} \mathcal{A}_{\mathrm{M}}^*} = (\ker \mathcal{A}_{\mathrm{M}})^\perp \subset \mathcal{X}_{\mathrm{M}} \perp \mathcal{X}_{\mathrm{E}} \supset (\ker \mathcal{A}_{\mathrm{E}})^\perp = \overline{\operatorname{ran} \mathcal{A}_{\mathrm{E}}^*}$. This complementarity of the orthogonal complements of the operator null spaces is essential for this proof. In order to prove the statement, we first calculate the adjoint operator $\mathcal{A}^*$ of the joint operator $\mathcal{A}$ in $\mathcal{X}$. It has to fulfil $\langle \mathcal{A}^* h, f \rangle_{\mathcal{X}} = \langle h, \mathcal{A} f \rangle_{\mathbb{R}^{\ell_\Sigma}}$ for all $f \in \mathcal{X}$ and $h \in \mathbb{R}^{\ell_\Sigma}$. The vector h can be decomposed into $h^{\mathsf{T}} = (h_{\mathrm{M}}^{\mathsf{T}}, h_{\mathrm{E}}^{\mathsf{T}})$ with $h_{\mathrm{M}} := \sum_{\ell=1}^{\ell_{\mathrm{M}}} (h)_\ell \varepsilon^\ell$ and $h_{\mathrm{E}} := \sum_{\ell=1}^{\ell_{\mathrm{E}}} (h)_{\ell_{\mathrm{M}}+\ell} \varepsilon^\ell$. Thus, via equation (6.1) we get

$$\begin{aligned}
\langle h, \mathcal{A} f \rangle_{\mathbb{R}^{\ell_\Sigma}} &= \langle h_{\mathrm{M}}, \mathcal{A}_{\mathrm{M}} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f \rangle_{\mathbb{R}^{\ell_{\mathrm{M}}}} + \langle h_{\mathrm{E}}, \mathcal{A}_{\mathrm{E}} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f \rangle_{\mathbb{R}^{\ell_{\mathrm{E}}}} \\
&= \langle \mathcal{A}_{\mathrm{M}}^* h_{\mathrm{M}}, \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f \rangle_{\mathcal{X}_{\mathrm{M}}} + \langle \mathcal{A}_{\mathrm{E}}^* h_{\mathrm{E}}, \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f \rangle_{\mathcal{X}_{\mathrm{E}}} \\
&= \langle \mathcal{A}_{\mathrm{M}}^* h_{\mathrm{M}}, \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f \rangle_{\mathcal{X}} + \langle \mathcal{A}_{\mathrm{E}}^* h_{\mathrm{E}}, \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f \rangle_{\mathcal{X}} \\
&= \langle \mathcal{A}_{\mathrm{M}}^* h_{\mathrm{M}}, f \rangle_{\mathcal{X}} + \langle \mathcal{A}_{\mathrm{E}}^* h_{\mathrm{E}}, f \rangle_{\mathcal{X}} \\
&= \langle \mathcal{A}_{\mathrm{M}}^* h_{\mathrm{M}} + \mathcal{A}_{\mathrm{E}}^* h_{\mathrm{E}}, f \rangle_{\mathcal{X}}.
\end{aligned}$$

The last-but-one step is valid due to the orthogonality and structure of the occurring spaces. Eventually, we obtain $\mathcal{A}^* h = \mathcal{A}_{\mathrm{M}}^* h_{\mathrm{M}} + \mathcal{A}_{\mathrm{E}}^* h_{\mathrm{E}}$ for all $h \in \mathbb{R}^{\ell_\Sigma}$, in particular for $h = g$. Setting $h = \mathcal{A} f$, we immediately arrive at

$$\mathcal{A}^* \mathcal{A} f = \mathcal{A}_{\mathrm{M}}^* \mathcal{A}_{\mathrm{M}} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f + \mathcal{A}_{\mathrm{E}}^* \mathcal{A}_{\mathrm{E}} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f. \tag{6.4}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We now utilize the fact that the theory for the RFMP was derived in general Hilbert spaces (see Theorem 4.2 and, e.g., [45, Sec. 5.4.1]). For this purpose, we define a new inner product on $\mathcal{X}$ as follows:

$$\langle f, \hat{f} \rangle_{\widetilde{\mathcal{X}}} := \langle \lambda_{\mathrm{E}}^{-1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f + \lambda_{\mathrm{M}}^{-1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f, \lambda_{\mathrm{E}}^{-1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} \hat{f} + \lambda_{\mathrm{M}}^{-1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} \hat{f} \rangle_{\mathcal{X}}.$$

Due to the orthogonality of $\mathcal{X}_{\mathrm{M}}$ and $\mathcal{X}_{\mathrm{E}}$, this inner product is well defined. Moreover, $\mathcal{X}_{\mathrm{M}}$ and $\mathcal{X}_{\mathrm{E}}$ remain orthogonal if $\langle \cdot, \cdot \rangle_{\widetilde{\mathcal{X}}}$ is used. The construction of the inner product yields

$$\begin{aligned} \lambda_{\mathrm{E}} \lambda_{\mathrm{M}} \|f\|_{\widetilde{\mathcal{X}}}^2 &= \langle \lambda_{\mathrm{M}}^{1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f + \lambda_{\mathrm{E}}^{1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f, \lambda_{\mathrm{M}}^{1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f + \lambda_{\mathrm{E}}^{1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f \rangle_{\mathcal{X}} \\ &= \lambda_{\mathrm{M}} \|\mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f\|_{\mathcal{X}_{\mathrm{M}}}^2 + \lambda_{\mathrm{E}} \|\mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f\|_{\mathcal{X}_{\mathrm{E}}}^2 . \end{aligned}$$

Hence, the Tikhonov functional for the previously known RFMP approach in the case of $(\mathcal{X}, \langle \cdot, \cdot \rangle_{\widetilde{\mathcal{X}}})$ as the domain of $\mathcal{A}$ and the regularization parameter $\lambda_{\mathrm{E}} \lambda_{\mathrm{M}}$ coincides with the modified Tikhonov functional (6.2) for the joint inversion. As we mentioned before, we can now apply the existing RFMP theory, which tells us that the RFMP for the joint inversion converges to f_∞ with

$$(\mathcal{A}_{\widetilde{\mathcal{X}}}^* \mathcal{A} + \lambda_{\mathrm{E}} \lambda_{\mathrm{M}} \mathcal{I}_{\widetilde{\mathcal{X}}}) f_\infty = \mathcal{A}_{\widetilde{\mathcal{X}}}^* g, \tag{6.5}$$

where $\mathcal{A}_{\widetilde{\mathcal{X}}}^*$ is the adjoint operator of $\mathcal{A}$ with respect to $\langle \cdot, \cdot \rangle_{\widetilde{\mathcal{X}}}$ and $\mathcal{I}_{\widetilde{\mathcal{X}}} = \mathcal{I}_{\mathcal{X}}$. To determine $\mathcal{A}_{\widetilde{\mathcal{X}}}^*$, we have a look at (with $h \in \mathbb{R}^{\ell_\Sigma}$)

$$\begin{aligned} \langle h, \mathcal{A} f \rangle_{\mathbb{R}^{\ell_\Sigma}} &= \langle \mathcal{A}^* h, f \rangle_{\mathcal{X}} \\ &= \langle \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} \mathcal{A}^* h + \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} \mathcal{A}^* h, \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f + \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f \rangle_{\mathcal{X}} \\ &= \langle \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} \mathcal{A}^* h, \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f \rangle_{\mathcal{X}} + \langle \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} \mathcal{A}^* h, \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f \rangle_{\mathcal{X}} \\ &= \langle \lambda_{\mathrm{E}}^{1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} \mathcal{A}^* h, \lambda_{\mathrm{E}}^{1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f \rangle_{\widetilde{\mathcal{X}}} + \langle \lambda_{\mathrm{M}}^{1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} \mathcal{A}^* h, \lambda_{\mathrm{M}}^{1/2} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f \rangle_{\widetilde{\mathcal{X}}} \\ &= \langle \lambda_{\mathrm{E}} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} \mathcal{A}^* h + \lambda_{\mathrm{M}} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} \mathcal{A}^* h, \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f + \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f \rangle_{\widetilde{\mathcal{X}}} \\ &= \langle \mathcal{A}_{\widetilde{\mathcal{X}}}^* h, f \rangle_{\widetilde{\mathcal{X}}}, \end{aligned}$$

where we used again the orthogonality of $\mathcal{X}_{\mathrm{M}}$ and $\mathcal{X}_{\mathrm{E}}$, and $\mathcal{A}^*$ is the adjoint operator of $\mathcal{A}$ with respect to $\langle \cdot, \cdot \rangle_{\mathcal{X}}$. Hence, $\mathcal{A}_{\widetilde{\mathcal{X}}}^* h = \lambda_{\mathrm{E}} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} \mathcal{A}^* h + \lambda_{\mathrm{M}} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} \mathcal{A}^* h$. Inserting this result into (6.5) and using (6.4), we obtain

$$\begin{aligned} &\lambda_{\mathrm{E}} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} \left(\mathcal{A}_{\mathrm{M}}^* \mathcal{A}_{\mathrm{M}} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f_\infty + \mathcal{A}_{\mathrm{E}}^* \mathcal{A}_{\mathrm{E}} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f_\infty \right) \\ &\quad + \lambda_{\mathrm{M}} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} \left(\mathcal{A}_{\mathrm{M}}^* \mathcal{A}_{\mathrm{M}} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f_\infty + \mathcal{A}_{\mathrm{E}}^* \mathcal{A}_{\mathrm{E}} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f_\infty \right) \\ &\quad + \lambda_{\mathrm{E}} \lambda_{\mathrm{M}} \left(\mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f_\infty + \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f_\infty \right) \\ &\qquad = \lambda_{\mathrm{E}} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} \mathcal{A}^* g + \lambda_{\mathrm{M}} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} \mathcal{A}^* g. \end{aligned}$$

Note that $\mathcal{P}_{\mathcal{X}_{\mathrm{M}}} \mathcal{A}_{\mathrm{E}}^* \mathcal{A}_{\mathrm{E}} \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f_\infty = 0$ and $\mathcal{P}_{\mathcal{X}_{\mathrm{E}}} \mathcal{A}_{\mathrm{M}}^* \mathcal{A}_{\mathrm{M}} \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f_\infty = 0$. By using once again the orthogonality of the occurring spaces and remembering that $\lambda_{\mathrm{E}} \lambda_{\mathrm{M}} > 0$, we now arrive at the following system of two equations:

$$\begin{aligned} \mathcal{A}_{\mathrm{M}}^* g_{\mathrm{M}} &= \left(\mathcal{A}_{\mathrm{M}}^* \mathcal{A}_{\mathrm{M}} + \lambda_{\mathrm{M}} \mathcal{I}_{\mathcal{X}_{\mathrm{M}}} \right) \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f_\infty, \quad \text{in } \mathcal{X}_{\mathrm{M}}, \\ \mathcal{A}_{\mathrm{E}}^* g_{\mathrm{E}} &= \left(\mathcal{A}_{\mathrm{E}}^* \mathcal{A}_{\mathrm{E}} + \lambda_{\mathrm{E}} \mathcal{I}_{\mathcal{X}_{\mathrm{E}}} \right) \mathcal{P}_{\mathcal{X}_{\mathrm{E}}} f_\infty, \quad \text{in } \mathcal{X}_{\mathrm{E}}. \end{aligned}$$

With equation (6.3), we get

$$(\mathcal{A}_{\mathrm{M}}^* \mathcal{A}_{\mathrm{M}} + \lambda_{\mathrm{M}} \mathcal{I}_{\mathcal{X}_{\mathrm{M}}}) f_{\mathrm{M},\infty} = \left(\mathcal{A}_{\mathrm{M}}^* \mathcal{A}_{\mathrm{M}} + \lambda_{\mathrm{M}} \mathcal{I}_{\mathcal{X}_{\mathrm{M}}} \right) \mathcal{P}_{\mathcal{X}_{\mathrm{M}}} f_\infty,$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$$(\mathcal{A}_{\mathrm{E}}^{*}\mathcal{A}_{\mathrm{E}} + \lambda_{\mathrm{E}}\mathcal{I}_{\mathcal{X}_{\mathrm{E}}})f_{\mathrm{E},\infty} = (\mathcal{A}_{\mathrm{E}}^{*}\mathcal{A}_{\mathrm{E}} + \lambda_{\mathrm{E}}\mathcal{I}_{\mathcal{X}_{\mathrm{E}}})\,\mathcal{P}_{\mathcal{X}_{\mathrm{E}}}f_{\infty}.$$

The two occurring operators are one-to-one since $\lambda_{\bullet} > 0$. Thus, $f_{\mathrm{M},\infty} = \mathcal{P}_{\mathcal{X}_{\mathrm{M}}}f_{\infty}$ and $f_{\mathrm{E},\infty} = \mathcal{P}_{\mathcal{X}_{\mathrm{E}}}f_{\infty}$. □

Since the solution obtained by the simultaneous inversion of both data sets equals the sum of the solutions obtained by the independent inversion of each single data set, no additional information is gained.

Furthermore, the separated inversion is more efficient than the simultaneous one for the following reasons. For the RFMP, the inner products between all dictionary elements are required; see equation (4.2). Thus, finding the maximum in equation (4.2) can be accelerated by splitting the inversion, since the memory capacity requirement for storing these inner products grows quadratically with the number of dictionary elements. Another problem in the simultaneous inversion of two different data sets emerges if the magnitudes of the data differ; see [15]. Therein, the data sets have to be weighted in order to obtain a good reconstruction. This is unnecessary in the case of a split inversion. Additionally, a split inversion allows a separate choice of the regularization parameters, i.e., it provides us with the possibility to fit each parameter to the particular problem, while a joint inversion requires the parallel choice of λ_{E} and λ_{M} within one single Tikhonov functional.

7. Implementation. For the implementation of the RFMP algorithm, several quantities are necessary, such as the forward operators, appropriate Hilbert spaces, a penalty term, a dictionary, and an initial approximation. The forward operators considered for the implementation are compoundly based on the Fredholm integral equations (5.1):

$$\mathcal{A}_{\mathrm{M}}J = \sum_{\ell=1}^{\ell_{\mathrm{M}}}(\mathcal{A}_{\mathrm{M}}^{\ell}J)\varepsilon^{\ell}, \quad \mathcal{A}_{\mathrm{E}}J = \sum_{\ell=\ell_{\mathrm{M}}+1}^{\ell_{\mathrm{E}}}(\mathcal{A}_{\mathrm{E}}^{\ell}J)\varepsilon^{\ell-\ell_{\mathrm{M}}}.$$

In our particular case, we have $\ell_{\mathrm{M}} = 102$ and $\ell_{\mathrm{E}} = 70$. For the implementation, we need to consider a fixed number of shells for the multiple-shell model, that is, $L = 3$. This is also known as the three-shell model. The used radii and conductivities are given by

$$\begin{aligned} &\varrho_0 = 0.071\,\mathrm{m}, && \varrho_1 = 0.072\,\mathrm{m}, && \varrho_2 = 0.079\,\mathrm{m}, && \varrho_3 = 0.085\,\mathrm{m},\\ &\sigma_0 = 0.330\,\mathrm{S\,m^{-1}}, && \sigma_1 = 1.000\,\mathrm{S\,m^{-1}}, && \sigma_2 = 0.042\,\mathrm{S\,m^{-1}}, && \sigma_3 = 0.330\,\mathrm{S\,m^{-1}}. \end{aligned}$$

Note that this model coincides with the one in [9, 19, 37] and is partially based on results stated in [27, Ch. 9.3]. For the sensor positions of the MEG device and the EEG cap, we have

$$\begin{aligned} &\|y_\ell\|_{\mathbb{R}^3} \in [1.4978\varrho_L, 2.0522\varrho_L], && \text{for all } \ell = 1,\dots,\ell_{\mathrm{M}},\\ &\|y_\ell\|_{\mathbb{R}^3} \in [\varrho_{L-1}, \varrho_L], && \text{for all } \ell = \ell_{\mathrm{M}}+1,\dots,\ell_{\mathrm{M}}+\ell_{\mathrm{E}}. \end{aligned}$$

The precise positions of the sensors in our synthetic test case as well as in the real data situation are visualized in Figure 7.1 and also used within the former numerical considerations; see [37]. Both sets of positions are irregularly distributed, with major gaps, for example in the region of the face or the lower half of the ball modelling the scalp.

Based on Algorithm 4.1, an initial approximation has to be chosen before starting the algorithm. Since we do not have a-priori knowledge of the neuronal current at hand, we choose the zero function to be the initial approximation.

7.1. The penalty term. For the domain, we choose $\mathrm{L}^2(B_{\varrho_0}, \mathbb{R}^3)$ as our largest space and restrict it to the directions of $\ker(\mathcal{A}_{\bullet})^{\perp}$, based on Theorem 4.2. Based on the results of Theorem 4.4, we can use a variety of norms induced by pseudo-differential operators for the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

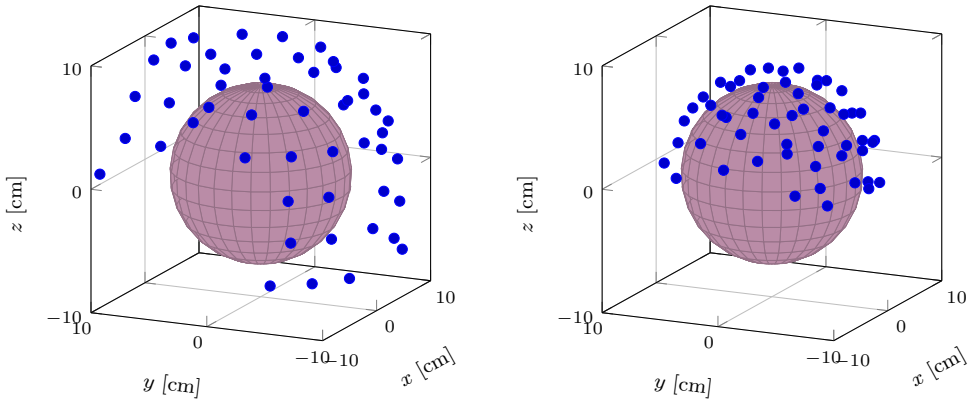

FIG. 7.1. *Positions of the MEG (left) and EEG (right) sensors around B_{ϱ_0} modelling the cerebrum.*

penalty term in order to control the smoothness of the results. For our numerical considerations, we choose several Sobolev norms for the penalty term based on the construction from Definition 4.3. The orthonormal basis functions in the domain required for the construction of the Sobolev spaces are chosen according to Theorem 3.1. More precisely, we use the following particular sequence depending on a parameter $s \in \mathbb{R}_0^+$, that is,

$$a_{m,n}^{(i)} := a_n^{(i)} \delta_{m,0} = \begin{cases} (n+\frac{1}{2})^s (n+\frac{1}{2})^s \delta_{m,0}, & \text{if } i = 2, \\ (n+\frac{1}{2})^s (n+\frac{3}{2})^s \delta_{m,0}, & \text{if } i = 3. \end{cases} \tag{7.1}$$

The corresponding Sobolev space will be denoted by $\mathcal{H}_s^{(i)}$. For the parameter s, we choose $s \in \{0, 1, 2\}$ within our numerical tests, where the case $s = 0$ corresponds to a classical L^2-regularization. Increasing the parameter s involves a faster decrease of the Fourier coefficients of the penalized function, which is somehow related to smoother functions.

7.2. The dictionaries. The elements of the dictionary have to be chosen in such a way that they are not in the null space of the forward operator $\mathcal{A}_\bullet$. For $N_\bullet = 5$, $L_\bullet \in \mathbb{N}$, and $I_\bullet := \{z_l \in B_{\varrho_0} \mid l = 1, \ldots, L_\bullet\}$, and a discrete set $H_\bullet \subset [0, 1)$, we choose the dictionary

$$\mathcal{D}_\bullet(N_\bullet, H_\bullet, I_\bullet) = \left\{ \tilde{g}_{0,n,j}^{(i)}(\varrho_0; \cdot) \right\}_{n=1,\ldots,N_\bullet,\ j=1,\ldots,2n+1} \cup \left\{ k_h^{(i)}(\cdot, z_l) \right\}_{h \in H_\bullet,\ l=1,\ldots,L_\bullet},$$

where $i = 2, 3$ depending on the particular problem. Here, the 35 used global orthonormal basis functions are given as in Theorem 3.1. Besides these functions, we use localized kernels of the form

$$\begin{aligned} k_h^{(i)}(x, z) &:= \sum_{\substack{(m,n) \in \mathbb{N}_0 \times \mathbb{N}_{0_i} \\ k_{m,n}^{(i)} \neq 0}} \sum_{j=1}^{2n+1} (k_{m,n}^{(i)})^{-2} \tilde{g}_{m,n,j}^{(i)}(\varrho_0; x) G_{m,n,j}^{(i)}(\varrho_0; z) \\ &= \frac{1}{4\pi\varrho_0^3} \begin{cases} \sum_{n=1}^{\infty} (2n+1)^2 h^n (rv/\varrho_0^2)^{n-1} \tilde{p}_n^{(2)}(\xi, \zeta), & \text{if } i = 2, \\ \sum_{n=1}^{\infty} (2n+3)(2n+1) h^n (rv/\varrho_0^2)^n \tilde{p}_n^{(3)}(\xi, \zeta), & \text{if } i = 3, \end{cases} \end{aligned} \tag{7.2}$$

with $x, z \in B_{\varrho_0}$ and the scalar-valued orthonormal basis from equation (3.3). In the series, we inserted the sequence $\{k_{m,n}^{(i)}\}_{(m,n) \in \mathbb{N}_0 \times \mathbb{N}_{0_i}}$, which is given by

$$k_{m,n}^{(i)} := h^{-n/2} \delta_{m,0}, \quad h \in H_\bullet := \{0.8, 0.9, 0.95, 0.99\}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 7.1
A maximal upper bound for the truncation error of the reproducing kernels and their forward solutions depending on the parameter $h \in (0,1)$ after $N = 230$ summands.

	MEG		EEG	
h	error $k_h^{(3)}$	error $\mathcal{A}_{\mathrm{M}} k_h^{(3)}$	error $k_h^{(2)}$	error $\mathcal{A}_{\mathrm{E}} k_h^{(2)}$
0.80	$1.9104 \cdot 10^{-25}$	$1.3107 \cdot 10^{-36}$	$8.6759 \cdot 10^{-26}$	$1.0741 \cdot 10^{-28}$
0.90	$1.8955 \cdot 10^{-13}$	$1.1575 \cdot 10^{-24}$	$8.6079 \cdot 10^{-14}$	$1.1719 \cdot 10^{-16}$
0.95	$6.7759 \cdot 10^{-8}$	$3.7216 \cdot 10^{-19}$	$3.0769 \cdot 10^{-8}$	$4.3234 \cdot 10^{-11}$
0.99	$1.2855 \cdot 10^{-3}$	$6.1445 \cdot 10^{-15}$	$5.8372 \cdot 10^{-4}$	$8.2779 \cdot 10^{-7}$

The addition theorem helps us to derive the kernel representation as a vector Legendre polynomial representation. Due to the particular construction of the orthonormal basis functions and the kernels, we immediately obtain that the dictionary is contained in $(\ker \mathcal{A}_\bullet)^\perp$. The centres z_l of the 50 620 kernels are chosen according to a modified Reuter grid on the ball presented in [29, 30] with a parameter $P = 8$; see Figure 7.2 (left).

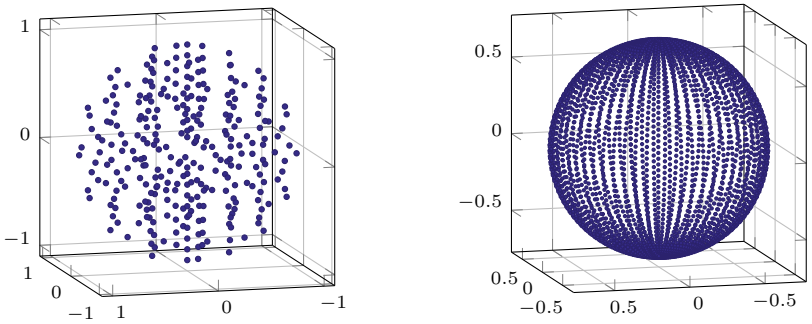

FIG. 7.2. *Modified Reuter grid on the ball for $P = 8$ with 289 grid points (left) and Driscoll–Healy grid with 3600 points (right).*

Note that the sequence of the kernel $\{k_{m,n}^{(i)}\}_{(m,n)\in\mathbb{N}_0\times\mathbb{N}_{0_i}}$ does not need to coincide with the sequence of the Sobolev space corresponding to our penalty term. Thus, within our dictionary, we can use reproducing kernels based on several different sequences. This is an advantage in comparison to a spline approach, where, for an interpolation or approximation, only one fixed sequence can be used.

For the implementation of the kernels, we need to truncate the occurring series. For the sake of computation time, we truncate all implemented series after 250 summands. Some estimates for the corresponding absolute maximal truncation errors depending on the parameter $h \in (0,1)$ are listed in Table 7.1. The truncation errors are estimated based on dominating geometric series for the kernel series representations and calculated via Mathematica [59]. The absolute truncation error seems large, but, due to an absolute value of the corresponding series around 3.5×10^3 (in the case of $k_h^{(3)}$ for $h = 0.99$), the relative error is of the order of 1×10^{-6}, which is still acceptable.

In some first numerical tests of the RFMP, a third type of dictionary elements was used. These functions were vector-valued radially invariant functions similar to the scalar-valued radially invariant functions used in [19] in order to solve a scalar inverse MEG problem. For $0 < a \le r \le b < \varrho_0$, the functions were of the form

$$\chi_{[a,b]}(r)\tilde{y}_{n,j}^{(i)}(\xi), \qquad n \in \mathbb{N}_{0_i},\ j = 1,\dots,2n+1.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

However, no dictionary element of this type was chosen in any numerical test. In [38], we showed that the harmonicity constraint in combination with the solenoidal condition is sufficient for a unique solution of the inverse problem. This unique solution coincides with the minimizer of the Tikhonov functional. In addition, the sequence obtained by the RFMP converges under certain conditions to the best approximate solution of the inverse problem; see Theorem 4.2. Based on the definition of the radially invariant functions, the harmonicity constraint $\Delta f_{n,j}^{(i)} = 0$ inside the entire ball B_{ϱ_0} is never fulfilled. This can be a possible reason for the RFMP not to choose these radially invariant functions.

The choice of the dictionary has an influence on the reconstruction, since we can only use a finite number of dictionary elements in the actual computation. Thus, the choice of the dictionary can also be understood as a regularization. For example, numerical tests showed that more orthonormal basis functions with a maximal degree higher than $N = 5$ yield poorer results, since the RFMP tries to reconstruct details and delicate structures with these high-degree orthonormal basis functions instead of using the localized reproducing kernels. This contradicts former numerical experiments in the geosciences. A reason for this behaviour may be the few data given in our particular problems, whereas, for example, 8500 data points were used in a synthetic downward continuation; see [57]. Thus, the few data points in our particular problem can also be interpolated with these basis functions. On the other hand, for the sake of computation time, we removed dictionary elements that were almost never chosen, such as reproducing kernels with a parameter h smaller than 0.8. However, these refinements are done by trial and error. In an ongoing project (e.g., [48, 54, 55]), a learning add-on (LIPMP) for the RFMP is being developed to avoid this brute-force method. Here, we use an infinite dictionary by modelling the minimization of the Tikhonov functional as a constrained non-linear optimization problem. The best basis of the LIPMPs can be used as a finite dictionary for the RFMP. The LIPMPs, until now, have only been applied to particular geomathematical inverse problems, where they yield promising results.

7.3. The preprocessing. In the case of the RFMP, the preprocessing is an excellent way to accelerate the algorithm, especially if a single dictionary is used for several RFMP runs. Quantities that are frequently used within the algorithm, such as inner products of the dictionary elements, and the application of the forward operator to the dictionary elements, are calculated for all dictionary elements in advance. Then, for a fixed regularization parameter, all quantities occurring in equation (4.2) are known. Thus, searching for the maximizer of equation (4.2) among all dictionary items reduces to finding the maximal entry of a vector whose length equals the number of dictionary elements, which can easily be parallelized. The same holds true for the ROFMP, since the additional backfitting consists of linear combinations of inner products of the dictionary elements.

For the preprocessing of the RFMP, we start with the calculation of the application of the operators $\mathcal{A}_{\mathrm{M}}$ and $\mathcal{A}_{\mathrm{E}}$, respectively, to the corresponding dictionary elements.

THEOREM 7.1. *The RFMP forward operator maps the orthogonal basis functions to*

$$(\mathcal{A}_{\mathrm{M}}\tilde{g}_{0,n,j}^{(3)}(\varrho_0;\cdot))(y) = -\mu_0\sqrt{\frac{n}{\varrho_0(2n+1)(2n+3)}}\left(\frac{\varrho_0}{s}\right)^{n+2}\tilde{y}_{n,j}^{(1)}(\eta),$$

$$(\mathcal{A}_{\mathrm{E}}\tilde{g}_{0,n,j}^{(2)}(\varrho_0;\cdot))(y) = \frac{1}{\sqrt{n}\varrho_0}\beta_n^{(L)}\left((n+1)\left(\frac{s}{\varrho_L}\right)^{2n+1}+n\right)\left(\frac{\varrho_0}{s}\right)^{n+1}Y_{n,j}(\eta),$$

where $n \in \mathbb{N}$ and $j = 1,\dots,2n+1$. Note that the sequence $(\beta_n^{(L)})_n$ is uniquely determined via the radii and the conductivity of the shells. A recursive representation can be found in [38]. For the reproducing kernels, we get, for all $h \in (0,1)$, $z \in B_{\varrho_0}$, and with the abbreviation

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$t \coloneqq \eta \cdot \zeta$, *the identities*

$$(\mathcal{A}_{\mathrm{M}} k_h^{(3)}(\cdot, z))(y) = -\frac{\mu_0}{4\pi} \sum_{n=1}^{\infty} \sqrt{\frac{n}{n+1}} \frac{(hv)^n}{s^{n+2}} \left((n+1)\eta P_n(t) - (\zeta - t\eta) P_n'(t)\right),$$

$$(\mathcal{A}_{\mathrm{E}} k_h^{(2)}(\cdot, z))(y) = \frac{1}{4\pi} \sum_{n=1}^{\infty} h^n \sqrt{\frac{(2n+1)^3}{n}} \beta_n^{(L)} \left((n+1)\left(\frac{s}{\varrho_L}\right)^{2n+1} + n\right) \frac{v^{n-1}}{s^{n+1}} P_n(t).$$

Proof. The proof can be found in Appendix A.1. ∎

We can calculate the inner products of the dictionary elements with each other in the preprocessing and store them efficiently in a vector containing a linearization of the lower triangular matrix.

THEOREM 7.2. *Let $i \in \{2, 3\}$ be fixed. Then we get, for all $n, \bar{n} \in \mathbb{N}$, $j = 1, \dots, 2n+1$, $\bar{j} = 1, \dots, 2\bar{n}+1$, $h, h' \in (0, 1)$, and $z_l, z_k \in B_{\varrho_0}$ with $l, k = 1, \dots, L_\bullet$, the following inner products between two dictionary functions:*

$$\left\langle \tilde{g}_{0,n,j}^{(i)}(\varrho_0; \cdot), \tilde{g}_{0,\bar{n},\bar{j}}^{(i)}(\varrho_0; \cdot) \right\rangle_{\mathcal{H}_\bullet^s} = (a_n^{(i)})^2 \delta_{n,\bar{n}} \delta_{j,\bar{j}},$$

$$\left\langle k_h^{(i)}(\cdot, z_l), \tilde{g}_{0,n,j}^{(i)}(\varrho_0; \cdot) \right\rangle_{\mathcal{H}_\bullet^s} = \sqrt{\frac{2t_n^{(i)} + 3}{\varrho_0^3}} (a_n^{(i)})^2 h^n \left(\frac{v_l}{\varrho_0}\right)^{t_n^{(i)}} Y_{n,j}(\zeta_l),$$

$$\left\langle k_h^{(i)}(\cdot, z_l), k_{h'}^{(i)}(\cdot, z_k) \right\rangle_{\mathcal{H}_\bullet^s}$$

$$= \frac{1}{4\pi\varrho_0^3} \sum_{\substack{n \in \mathbb{N} \\ a_n^{(i)} \neq 0}} (a_n^{(i)})^2 (hh')^n (2n+1)(2t_n^{(i)} + 3) \left(\frac{v_l v_k}{\varrho_0^2}\right)^{t_n^{(i)}} P_n(t),$$

where $t \coloneqq \zeta_l \cdot \zeta_k \in [-1, 1]$. Recall that the Sobolev space $\mathcal{H}_\bullet^s$ is defined with respect to the sequence $\{a_{m,n}^{(i)}\}_{m \in \mathbb{N}_0,\, n \in \mathbb{N}_{0_i}}$ in equation (7.1).

Proof. The proof can be found in Appendix A.2. ∎

7.4. The visualization. For a visualization of the numerical results, the approximations are plotted on a spherical point grid within the cerebrum B_{ϱ_0}. The radius of this sphere is typically chosen as $0.95\varrho_0$ and the angle of the plot is chosen such that the upper hemisphere is seen. More precisely, we use the equiangular Driscoll–Healy grid introduced in [11]. The resulting grid with 3600 points is plotted in Figure 7.2 (right).

7.5. Synthetic test current. In order to validate the results obtained by the reconstruction methods, we construct a synthetic test case as realistically as possible similar to the ones in [19, 37] separately for both applications. In each case, the corresponding synthetic test current is assumed to be in the orthogonal complement of the operator null space in order to handle the non-uniqueness of the ill-posed problems. This implies that the synthetic current is harmonic and solenoidal.

The current is based on the classical Abel–Poisson kernel, which is also used in [19] for building an appropriate test case:

$$\frac{1}{4\pi} \frac{1 - h^2}{(1 + h^2 - 2ht)^{3/2}} = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} h^n P_n(t), \quad h \in (-1, 1),\ t \in [-1, 1]. \tag{7.3}$$

The Abel–Poisson kernel is combined with spherical differential operators in order to achieve a test current that is entirely located in the orthogonal complement of the operator null space.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

DEFINITION 7.3. *Let $h \in [0,1)$ be fixed and $z \in B_{\varrho_0}$ be given. Then the synthetic MEG current I_{M} and the synthetic EEG current I_{E} are, for all $x \in B_{\varrho_0}$, defined by*

$$I_{\mathrm{M}}(x;h,z) := L_\xi^*\left(\sum_{n=1}^{\infty} \frac{2n+1}{4\pi}\left(\frac{hrv}{\varrho_0^2}\right)^n P_n(\xi\cdot\zeta)\right),$$
$$I_{\mathrm{E}}(x;h,z) := \nabla_x\left(\sum_{n=1}^{\infty} \frac{2n+1}{4\pi}\left(\left(\frac{hrv}{\varrho_0^2}\right)^n P_n(\xi\cdot\zeta)\right)\right).$$

In a first step, we need to verify that the synthetic currents do not have parts in the null spaces of the respective forward operators.

THEOREM 7.4. *Let the synthetic currents be defined as in Definition 7.3. Then, for all $h \in [0,1)$ and all $z \in B_{\varrho_0}$, we have $I_\bullet(\cdot;h,z) \in (\ker \mathcal{A}_\bullet)^\perp$.*

Proof. The proof can be found in Appendix B.1. ∎

For the implementation of the synthetic current, the series representation is not appropriate since the required truncation of the series will always result in approximation errors. Hence, we calculate the closed representations of the synthetic currents.

LEMMA 7.5. *Let the synthetic current $I_\bullet$ be defined as in Definition 7.3 and abbreviations be given by $p := \xi\cdot\zeta$ and $q := q(r) := hrv/\varrho_0^2$ for all $h \in [0,1)$ and $z \in B_{\varrho_0}$. Then,*

$$I_{\mathrm{M}}(x;h,z) = \frac{3}{4\pi}\frac{(1-q^2)q}{(1+q^2-2q(\xi\cdot\zeta))^{5/2}}(\xi\wedge\zeta), \qquad x \in B_{\varrho_0},$$
$$I_{\mathrm{E}}(x;h,z) = \frac{1}{4\pi}\frac{q^2(-5+q^2+4qp)\xi + 3(1-q^2)q\zeta}{r(1+q^2-2qp)^{5/2}}, \quad x \in B_{\varrho_0}.$$

Proof. The proof can be found in Appendix B.2. ∎

Examples of a linear combination of two synthetic currents for the inverse MEG as well as for the inverse EEG problem are shown in Figures 8.2 (left) and 8.4 (left). In both cases, the occurring parameters are chosen according to Table 7.2, where κ_l denotes the coefficients.

TABLE 7.2
Parameters for the synthetic test current for the inverse MEG and EEG problem.

l	κ_l	h_l	$\lvert z_l\rvert$	ζ_l
1	1	0.9	$0.85\varrho_0$	$(0,-2,1)^\intercal/\sqrt{5}$
2	1.5	0.8	$0.9\varrho_0$	$(-1,1,1)^\intercal/\sqrt{3}$

7.6. Synthetic data. To calculate the corresponding synthetic data, we need to apply the MEG and EEG operators $\mathcal{A}_{\mathrm{M}}$ and $\mathcal{A}_{\mathrm{E}}$ introduced in Definition 6.1 to the synthetic currents from Definition 7.3.

THEOREM 7.6. *Let $y \in \overline{B_{\varrho_L}^{\mathrm{ext}}}$ in the MEG case and let $y \in S_{[\varrho_{L-1},\varrho_L]}$ in the EEG case. Then we get*

$$\left(\mathcal{A}_{\mathrm{M}}J_{\mathrm{M}}\right)(y) = -\frac{\mu_0}{4\pi}\varrho_0^3\sum_{l=1}^{2}\kappa_l\sum_{n=1}^{\infty}\sqrt{\frac{n^2(n+1)(2n+1)}{(2n+3)^2}}\frac{(h_l v_l)^n}{s^{n+2}}\tilde{p}_n^{(1)}(\eta;\zeta_l),$$
$$\left(\mathcal{A}_{\mathrm{E}}J_{\mathrm{E}}\right)(y) = \frac{\varrho_0}{4\pi}\sum_{l=1}^{2}\kappa_l\sum_{n=1}^{\infty}\beta_n^{(L)}\frac{(h_l v_l)^n}{s^{n+1}}(2n+1)\left((n+1)\left(\frac{s}{\varrho_L}\right)^{2n+1}+n\right)P_n(\zeta_l\cdot\eta)$$

with $\kappa_l \in \mathbb{R}$, $h_l \in [0,1)$, and $z_l \in B_{\varrho_0}$ for $l = 1,2$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proof. The proof can be found in Appendix B.3. ∎

Unfortunately, we are not able to find a closed representation of either $\mathcal{A}_{\mathrm{M}}(I_{\mathrm{M}}(\cdot;h,z))$ or $\mathcal{A}_{\mathrm{E}}(I_{\mathrm{E}}(\cdot;h,z))$. Therefore, we need to truncate the series for the computation of the synthetic data. Thus, we are interested in a bound for the truncation error.

LEMMA 7.7. *For all $h \in [0,1)$, $z \in B_{\varrho_0}$, $y \in \overline{B}^{\mathrm{ext}}_{\varrho_L}$, and $N \in \mathbb{N}$, the following estimate holds true:*

$$\varrho_0^3 \left| \sum_{n=N+1}^{\infty} \sqrt{\frac{n^2(n+1)(2n+1)}{(2n+3)^2}} \frac{(hv)^n}{s^{n+2}} \tilde{p}_n^{(1)}(\eta;\zeta) \right| \leq \frac{\varrho_0^3}{\varrho_L^2} \frac{(h\varrho_0/\varrho_L)^{N+1}(N+2)^2}{(1-h\varrho_0/\varrho_L)^3}.$$

In addition, for all $h \in [0,1)$, $z \in B_{\varrho_0}$, $y \in S_{[\varrho_{L-1},\varrho_L]}$, and $N \in \mathbb{N}$, we get

$$\begin{aligned} &\varrho_0 \left| \sum_{n=N+1}^{\infty} \beta_n^{(L)} \frac{h^n v^n}{s^{n+1}} (2n+1) \left((n+1) \left(\frac{s}{\varrho_L} \right)^{2n+1} + n \right) P_n(\zeta \cdot \eta) \right| \\ &\qquad \lesssim \frac{9\varrho_0}{2\sigma_L \varrho_{L-1} C^{(L)}} \frac{(h\varrho_0/\varrho_{L-1})^{N+1}(N+1)}{(1-h\varrho_0/\varrho_{L-1})^2}, \end{aligned}$$

where $C^{(L)}$ is part of an upper bound for the sequence $\{n\beta_n^{(L)}\}_{n\in\mathbb{N}}$; see [38, Lem. 4.2] for more information about its asymptotic behaviour. Note that the consequently required condition on $\{n\beta_n^{(L)}\}_{n\in\mathbb{N}}$ is fulfilled in our particular numerical setting.

Proof. The proof can be found in Appendix B.4. ∎

If a maximal truncation error, for example the machine precision, is desired, then the results stated in Table 7.3 yield the required number of summands depending on the parameter h. According to the results from Table 7.3, which were also obtained via Mathematica [59], we choose $N_\bullet = 250$, to guarantee a small error over all series truncations.

In addition, numerical summation always implies a numerical error due to the finite precision of the floating-point numbers. In order to reduce the truncation and round-off error incurred by the summation, we use the Kahan summation algorithm (see [28]) for the MEG forward operator. From the results stated in Table 7.3, we can deduce that $N_{\mathrm{M}} \times \mathrm{eps} \leq 1$, where eps is the machine precision. Thus, the error of compensated summation is effectively $\mathcal{O}(\mathrm{eps})$, which is independent of N; see [28]. For the calculation of the electric potential, we use the Clenshaw algorithm for the summation of Legendre polynomials; see [8] or [43, p. 48].

TABLE 7.3
Number of required summands $N_\bullet$ for a uniform maximal truncation error of the synthetic test data depending on h; the largest truncation parameter is written in bold.

	$h = 0.8$		$h = 0.9$	
error	N_{M}	N_{E}	N_{M}	N_{E}
1×10^{-4}	48	62	74	103
1×10^{-5}	54	69	83	114
1×10^{-16}	121	149	**177**	**237**

Besides non-noisy data, we want to generate noisy data for the synthetic tests by adding white Gaussian noise. In this case, let $\{g_i\}_{1\leq i\leq N}$ be the data. Then the additive noise for each component g_i of the data vector is normally distributed with zero mean and standard deviation $\delta|g_i|$, where $\delta \in [0, 0.1]$ is the (relative) noise level.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

8. Numerical results. Within our synthetic test case, we used the RFMP as well as the ROFMP for the inversion. The RFMP is stopped after a fixed number of 600 iterations. The maximal number of iterations for the ROFMP is 250, where the algorithm is restarted after $K = 25$ steps. However, due to the fast decrease of the relative residual obtained by the ROFMP, we additionally stop the algorithm if the relative residual after a restart is less than or equal to $2\,\%$.

To judge the quality of the reconstruction obtained by the R(O)FMP, we consider the approximation error measured via the normalized root-mean-square difference (NRMSE) from the reconstruction to the test current on the 3600-points Driscoll–Healy grid in Figure 7.2 (right). In addition, we analyse the data misfit by the residual normalized with respect to the data $g^{\delta\bullet}$. Its typical evolution during the synthetic data inversion via the R(O)FMP is shown in Figure 8.1 in the case of $\delta_\bullet = 0.05$. In the EEG case, the relative residual drops below $5\,\%$ after six (RFMP), respectively four (ROFMP), iterations, which is desired for the noise level of $5\,\%$ according to the discrepancy principle; see [1, 2, 14]. Although the curve of the relative residual in the MEG case is qualitatively similar to the curve in the EEG case, the decay is significantly slower. One possible reason for this behaviour is the severe ill-posedness of the problems and the faster decay of the MEG singular values.

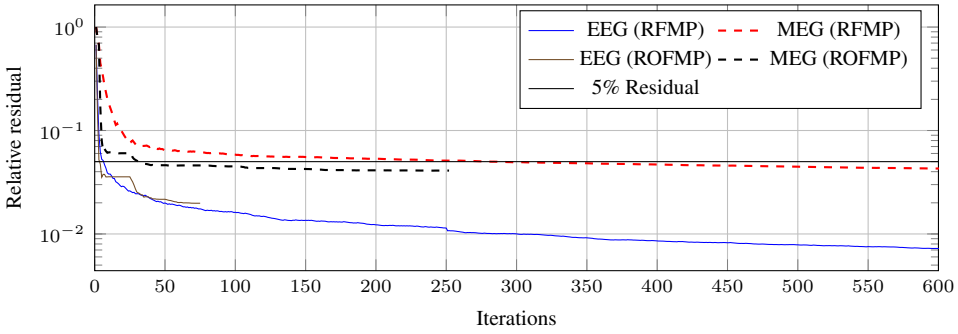

FIG. 8.1. *Evolution of the relative residual during the R(O)FMP iterations for* $\delta_\bullet = 0.05$.

According to the implemented parameter choice methods, the R(O)FMP is started with up to 300 different regularization parameters λ. First, we try to find the *optimal* regularization parameter λ_* that minimizes the NRMSE. Corresponding to this regularization parameter, the values of the approximation norm, the NRMSE, and the relative residual obtained via the R(O)FMP are listed for all EEG test cases in Table 8.1. Since the inverse MEG and EEG problems are severely ill-posed problems, it is not surprising that the approximation error is larger than the relative residual and the noise level throughout almost all numerical tests.

Based on Table 8.1 for the EEG case, the ROFMP yields more accurate results than the RFMP, except for the largest noise levels. For the MEG case, a comparable table is given in Table 8.2. The obtained NRMSE reduces by an order of magnitude compared to the RFMP in the non-noisy case. This effect is also observed for the other noise levels but is much less pronounced, as the difference shrinks with increasing noise level. In addition, the RFMP produces some artefacts in the case of higher noise levels, whereas the reconstruction via the ROFMP is more accurate; see Figure 8.2. The ROFMP has the additional advantage that it needs a smaller number of dictionary elements for the reconstruction (sparsity of the solution); see Table 8.3.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 8.1
Comparison of NRMSE and relative residual of RFMP and ROFMP reconstruction for the EEG inversion.

		RFMP			ROFMP		
δ_{E}	s	$\left\|J_{\lambda_*}^{\delta_{\mathrm{E}}}\right\|^2_{\mathcal{H}_s^{(2)}}$	NRMSE	rel. residual	$\left\|J_{\lambda_*}^{\delta_{\mathrm{E}}}\right\|^2_{\mathcal{H}_s^{(2)}}$	NRMSE	rel. residual
0.00	0	$2.011 \cdot 10^{0}$	$2.490 \cdot 10^{-2}$	$3.015 \cdot 10^{-3}$	$7.967 \cdot 10^{-3}$	$6.357 \cdot 10^{-3}$	$5.655 \cdot 10^{-3}$
0.00	1	$4.477 \cdot 10^{1}$	$1.597 \cdot 10^{-2}$	$4.980 \cdot 10^{-3}$	$5.915 \cdot 10^{1}$	$3.647 \cdot 10^{-3}$	$1.472 \cdot 10^{-3}$
0.00	2	$4.179 \cdot 10^{2}$	$1.210 \cdot 10^{-2}$	$2.394 \cdot 10^{-3}$	$5.769 \cdot 10^{3}$	$2.400 \cdot 10^{-3}$	$3.891 \cdot 10^{-3}$
0.01	0	$2.025 \cdot 10^{0}$	$2.638 \cdot 10^{-2}$	$3.037 \cdot 10^{-3}$	$1.095 \cdot 10^{-2}$	$1.252 \cdot 10^{-2}$	$7.306 \cdot 10^{-3}$
0.01	1	$6.954 \cdot 10^{1}$	$1.681 \cdot 10^{-2}$	$5.944 \cdot 10^{-3}$	$1.963 \cdot 10^{1}$	$6.456 \cdot 10^{-3}$	$9.892 \cdot 10^{-3}$
0.01	2	$9.064 \cdot 10^{3}$	$1.229 \cdot 10^{-2}$	$6.729 \cdot 10^{-3}$	$4.355 \cdot 10^{2}$	$6.517 \cdot 10^{-3}$	$4.865 \cdot 10^{-3}$
0.05	0	$1.399 \cdot 10^{0}$	$3.413 \cdot 10^{-2}$	$9.911 \cdot 10^{-3}$	$1.233 \cdot 10^{-1}$	$2.599 \cdot 10^{-2}$	$1.719 \cdot 10^{-2}$
0.05	1	$8.751 \cdot 10^{1}$	$2.298 \cdot 10^{-2}$	$1.136 \cdot 10^{-2}$	$1.454 \cdot 10^{1}$	$1.376 \cdot 10^{-2}$	$1.982 \cdot 10^{-2}$
0.05	2	$7.092 \cdot 10^{3}$	$1.974 \cdot 10^{-2}$	$1.494 \cdot 10^{-2}$	$5.963 \cdot 10^{3}$	$1.611 \cdot 10^{-2}$	$1.755 \cdot 10^{-2}$
0.10	0	$4.286 \cdot 10^{-1}$	$5.752 \cdot 10^{-2}$	$4.437 \cdot 10^{-2}$	$1.383 \cdot 10^{-1}$	$7.741 \cdot 10^{-2}$	$1.841 \cdot 10^{-2}$
0.10	1	$4.286 \cdot 10^{1}$	$3.927 \cdot 10^{-2}$	$6.155 \cdot 10^{-2}$	$2.079 \cdot 10^{1}$	$4.700 \cdot 10^{-2}$	$3.866 \cdot 10^{-2}$
0.10	2	$9.183 \cdot 10^{3}$	$4.578 \cdot 10^{-2}$	$4.986 \cdot 10^{-2}$	$2.055 \cdot 10^{2}$	$4.522 \cdot 10^{-2}$	$5.949 \cdot 10^{-2}$

TABLE 8.2
Comparison of NRMSE and relative residual of RFMP and ROFMP reconstruction for the MEG inversion.

		RFMP			ROFMP		
δ_{M}	s	$\left\|J_{\lambda_*}^{\delta_{\mathrm{M}}}\right\|^2_{\mathcal{H}_s^{(3)}}$	NRMSE	rel. residual	$\left\|J_{\lambda_*}^{\delta_{\mathrm{M}}}\right\|^2_{\mathcal{H}_s^{(3)}}$	NRMSE	rel. residual
0.00	0	$8.627 \cdot 10^{-1}$	$6.692 \cdot 10^{-2}$	$1.779 \cdot 10^{-2}$	$9.077 \cdot 10^{-1}$	$1.456 \cdot 10^{-2}$	$2.304 \cdot 10^{-3}$
0.00	1	$5.200 \cdot 10^{1}$	$6.489 \cdot 10^{-2}$	$1.652 \cdot 10^{-2}$	$6.335 \cdot 10^{1}$	$8.414 \cdot 10^{-3}$	$3.575 \cdot 10^{-3}$
0.00	2	$5.155 \cdot 10^{4}$	$8.539 \cdot 10^{-2}$	$1.805 \cdot 10^{-2}$	$1.868 \cdot 10^{4}$	$6.477 \cdot 10^{-3}$	$3.460 \cdot 10^{-3}$
0.01	0	$8.854 \cdot 10^{-1}$	$7.442 \cdot 10^{-2}$	$1.952 \cdot 10^{-2}$	$9.181 \cdot 10^{-1}$	$2.639 \cdot 10^{-2}$	$8.617 \cdot 10^{-3}$
0.01	1	$5.066 \cdot 10^{1}$	$6.232 \cdot 10^{-2}$	$1.781 \cdot 10^{-2}$	$5.993 \cdot 10^{1}$	$1.308 \cdot 10^{-2}$	$1.553 \cdot 10^{-2}$
0.01	2	$2.157 \cdot 10^{4}$	$7.010 \cdot 10^{-2}$	$1.489 \cdot 10^{-2}$	$1.421 \cdot 10^{4}$	$1.080 \cdot 10^{-2}$	$1.283 \cdot 10^{-2}$
0.05	0	$8.823 \cdot 10^{-1}$	$8.545 \cdot 10^{-2}$	$4.018 \cdot 10^{-2}$	$8.084 \cdot 10^{-1}$	$9.104 \cdot 10^{-2}$	$4.027 \cdot 10^{-2}$
0.05	1	$4.780 \cdot 10^{1}$	$6.862 \cdot 10^{-2}$	$4.289 \cdot 10^{-2}$	$3.572 \cdot 10^{1}$	$6.554 \cdot 10^{-2}$	$4.098 \cdot 10^{-2}$
0.05	2	$2.209 \cdot 10^{4}$	$7.537 \cdot 10^{-2}$	$4.114 \cdot 10^{-2}$	$6.666 \cdot 10^{3}$	$4.696 \cdot 10^{-2}$	$3.749 \cdot 10^{-2}$
0.10	0	$8.176 \cdot 10^{-1}$	$1.030 \cdot 10^{-1}$	$7.667 \cdot 10^{-2}$	$8.096 \cdot 10^{-1}$	$1.072 \cdot 10^{-1}$	$6.962 \cdot 10^{-2}$
0.10	1	$3.724 \cdot 10^{1}$	$7.981 \cdot 10^{-2}$	$7.545 \cdot 10^{-2}$	$4.923 \cdot 10^{1}$	$7.699 \cdot 10^{-2}$	$6.013 \cdot 10^{-2}$
0.10	2	$4.596 \cdot 10^{3}$	$6.578 \cdot 10^{-2}$	$7.488 \cdot 10^{-2}$	$5.044 \cdot 10^{3}$	$6.043 \cdot 10^{-2}$	$7.638 \cdot 10^{-2}$

In addition, we can conclude that, in the MEG and EEG synthetic tests, the L^2-regularization ($s = 0$) yields the highest approximation errors for all noise levels. The ROFMP with $s = 2$ produces the best results among the tested cases for the reconstruction of the synthetic test current J_{M} (MEG). In the EEG case, the data listed in Table 8.1 is not so expressive. Both regularization terms (corresponding to $s = 1$ and $s = 2$) yield good results; see Figure 8.3. Using the RFMP, the smallest NRMSE throughout the noise levels $\delta_{\mathrm{E}} \in \{0, 0.01, 0.05\}$ is obtained for $s = 2$. However, with $10\,\%$ noise on the data, the regularization term corresponding to $s = 1$ yields a slightly smaller error. If the ROFMP is used, $s = 2$ yields better results on non-noisy data and for $\delta_{\mathrm{E}} = 0.1$, whereas $s = 1$ yields smaller NRMSEs for the other noise levels. Figure 8.4 shows that the reconstruction of the synthetic current is still good even for $10\,\%$ noise for such a severely ill-posed problem. The active regions are easily identifiable and the reconstruction is not too blurry. In conclusion, the regularization terms with $s \in \{1, 2\}$ should be preferred over $s = 0$ in the EEG case, since we cannot find a clear winner.

Combining all these numerical results, we eventually conclude that the ROFMP with regularization term for $s = 2$ generates good reconstructions for the synthetic MEG test case

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 8.3
Number of ROFMP iterations depending on noise levels and regularization terms.

	$\delta_\bullet = 0$			$\delta_\bullet = 0.01$			$\delta_\bullet = 0.05$			$\delta_\bullet = 0.1$		
Iterations	$\mathcal{H}_0^{(i)}$	$\mathcal{H}_1^{(i)}$	$\mathcal{H}_2^{(i)}$	$\mathcal{H}_0^{(i)}$	$\mathcal{H}_1^{(i)}$	$\mathcal{H}_2^{(i)}$	$\mathcal{H}_0^{(i)}$	$\mathcal{H}_1^{(i)}$	$\mathcal{H}_2^{(i)}$	$\mathcal{H}_0^{(i)}$	$\mathcal{H}_1^{(i)}$	$\mathcal{H}_2^{(i)}$
EEG	25	25	25	25	25	25	25	75	50	75	250	250
MEG	25	25	25	25	25	25	250	250	250	250	250	250

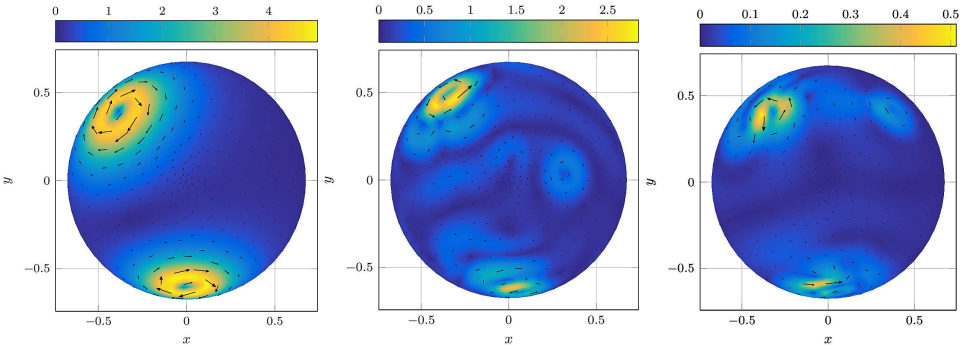

FIG. 8.2. *Comparison of RFMP (middle) and ROFMP (right) reconstruction with 1 % noise (MEG) to the exact solution (left). The images in the middle and on the right show the absolute approximation errors.*

with and without noisy data. Therefore, we choose this regularization term for the inversion of real magnetic flux data. For the inversion of the real electric potential data, we will additionally take the regularization term corresponding to $s = 1$ into account.

8.1. Performance benchmark. Although the ROFMP requires essentially fewer iterations than the RFMP, it is, with few exceptions, the slower algorithm due to the required backfitting steps. The CPU time is visualized in Figure 8.5 in the form of a boxplot. For each test case and parameter combination, the R(O)FMP is started with 100 different regularization parameters and the required CPU time is recorded. The box contains timings that are between 25 % and 75 % of all 100 observed values. The horizontal lines inside the boxes mark the median of the CPU times. Finally, the very first and last horizontal lines, represent the minimum and maximum of the measured timings. In addition, the time required for the regularization parameter with minimal NRMSE is marked with a cross. The boxes corresponding to the RFMP are significantly smaller than those belonging to the ROFMP. The reason for this behaviour is that the ROFMP stops depending on the discrepancy principle, which causes inconsistent timings. However, due to the ROFMP approximations being significantly sparser than the RFMP ones, the postprocessing accelerates. For example, the time required for plotting the approximation depends almost linearly on the number of chosen dictionary elements.

8.2. Parameter choice methods. In a real data situation, we cannot determine the *optimal* regularization parameter via the NRMSE, as an exact reference solution is not available. Since the choice of the regularization parameter is essential for the quality of the result obtained by the R(O)FMP, we compare several parameter choice methods using the synthetic data in order to select suitable methods for the problem at hand. We implemented the L-curve method (LCM) and an automatic version (LCA) of it. In addition, we used the discrepancy principle (DP), generalized cross-validation (GCV), the strong robust generalized cross-validation (SRGCV), a modified generalized cross-validation (MGCV), the residual method (RM), and the quasi-optimality criterion (QOC). All these methods have been used before in the context

ETNA
Kent State University and
Johann Radon Institute (RICAM)

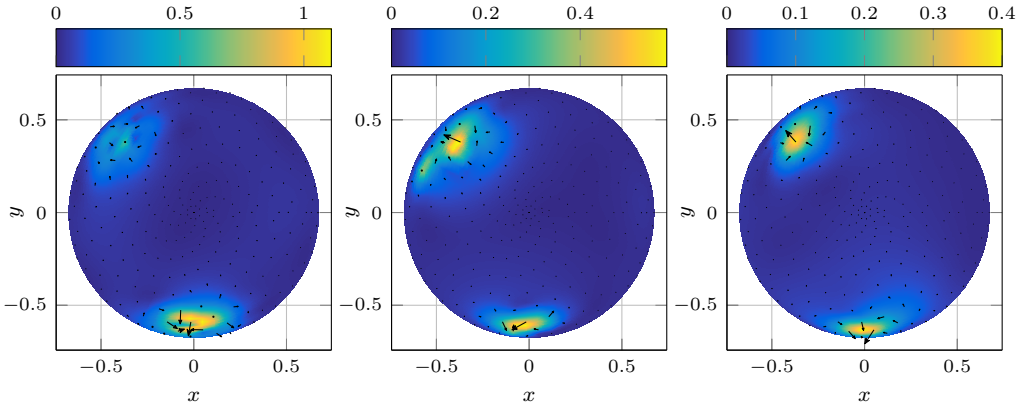

FIG. 8.3. *Deviation of the reconstruction J_{λ_*} by the ROFMP to the exact solution (non-noisy case) depending on the penalty term $\mathcal{H}_s^{(2)}$ for $s = 0$ (left), $s = 1$ (middle), and $s = 2$ (right) for the EEG case.*

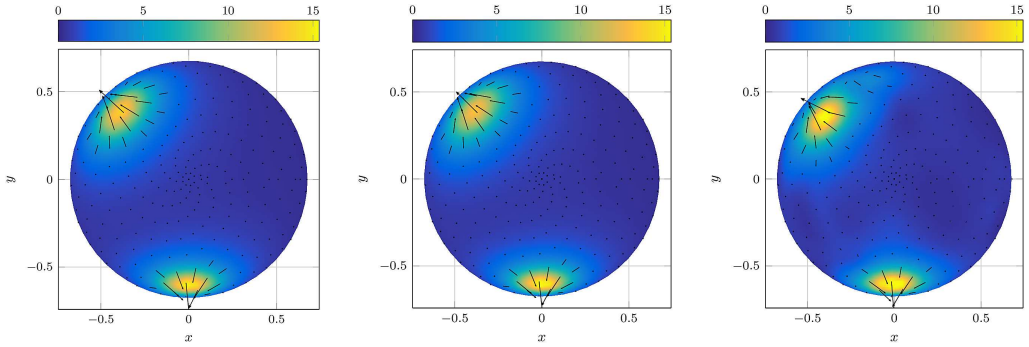

FIG. 8.4. *Exact solution (left) and approximation $J_{\lambda_*}^{\delta_{\mathrm{E}}}$ of the neuronal current (EEG) via the ROFMP in $\mathcal{H}_1^{(2)}$ from 1 % noise (middle) and 10 % noise (right).*

of regularization of ill-posed problems, especially in geomathematics and in combination with the RFMP; see [1, 2, 25].

Table 8.4 reveals that the MGCV and the LCA (see Figure 8.5 (right)) yield the best results for noise levels below $5\,\%$. However, these methods fail for the inversion of $5\,\%$-noisy data. The remaining parameter choice methods yield better but not good results for this noise level, since the corresponding NRMSEs are 2–3 times as high as with the optimal parameter. The GCV and the (strong) robust GCV yield the smallest NRMSEs among all parameter choice methods for $\delta_{\mathrm{M}} = 0.1$, but the error is twice as high as with the optimal parameter. The evaluation of the parameter choice methods for the EEG problem via the ROFMP with $s = 1$ reveals that the LCM is the parameter choice method with the best approximation error except for non-noisy data, where the MGCV yields a slightly smaller NRMSE; see Table 8.5. This can also be observed in the case $s = 2$. Based on the observations and the results in Table 8.4, we conclude that how to find the best or even a good regularization parameter for the R(O)FMP combined with the inverse MEG and EEG problem depends on the noise level and is still an open question. In [25], first experiments regarding the choice of λ for the RFMP were performed, where the L-curve method turned out to be a good option.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 8.4
NRMSE and relative residual achieved by $J_{\lambda_*}^{\delta_{\mathrm{M}}}$ (MEG) via the ROFMP with $s = 2$, where λ_* is chosen according to different parameter choice methods.

Parameter choice method		$\delta_{\mathrm{M}} = 0$	$\delta_{\mathrm{M}} = 0.01$	$\delta_{\mathrm{M}} = 0.05$	$\delta_{\mathrm{M}} = 0.1$
NRMSE	NRMSE	0.006 47	0.010 79	0.046 95	0.060 43
	Rel. residual	0.003 45	0.012 83	0.037 49	0.076 37
GCV	NRMSE	0.111 72	0.113 77	0.110 40	0.120 78
	Rel. residual	0.043 08	0.041 08	0.070 69	0.103 40
MGCV	NRMSE	0.010 89	0.036 35	0.403 17	0.427 26
	Rel. residual	0.001 36	0.005 35	0.017 78	0.026 78
RGCV	NRMSE	0.111 72	0.113 77	0.110 40	0.120 78
	Rel. residual	0.043 08	0.041 08	0.070 69	0.103 40
SRGCV	NRMSE	0.111 72	0.113 77	0.110 40	0.120 78
	Rel. residual	0.043 08	0.041 08	0.070 69	0.103 40
RM	NRMSE	0.111 72	0.113 77	0.110 40	0.120 78
	Rel. residual	0.043 08	0.041 08	0.070 69	0.103 40
LCA	NRMSE	0.010 89	0.036 35	0.433 89	0.495 96
	Rel. residual	0.001 36	0.005 35	0.018 07	0.027 82
LCM	NRMSE	0.024 06	0.063 90	0.078 57	0.209 37
	Rel. residual	0.006 10	0.020 12	0.031 34	0.040 74

TABLE 8.5
NRMSE and relative residual achieved by $J_{\lambda_*}^{\delta_{\mathrm{E}}}$ (EEG) via the ROFMP with $s = 1$, where λ_* is chosen according to different parameter choice methods.

Parameter choice method		$\delta_{\mathrm{E}} = 0$	$\delta_{\mathrm{E}} = 0.01$	$\delta_{\mathrm{E}} = 0.05$	$\delta_{\mathrm{E}} = 0.1$
NRMSE	NRMSE	0.003 64	0.006 45	0.013 75	0.050 41
	Rel. residual	0.001 47	0.009 89	0.019 81	0.041 51
GCV	NRMSE	0.042 58	0.057 26	0.052 32	0.072 23
	Rel. residual	0.016 91	0.018 64	0.019 46	0.048 45
MGCV	NRMSE	0.005 96	0.012 84	0.046 64	0.111 27
	Rel. residual	0.001 35	0.002 60	0.007 47	0.011 34
RGCV	NRMSE	0.042 58	0.057 26	0.052 32	0.072 23
	Rel. residual	0.016 91	0.018 64	0.019 46	0.048 45
SRGCV	NRMSE	0.042 58	0.057 26	0.052 32	0.072 23
	Rel. residual	0.016 91	0.018 64	0.019 46	0.048 45
RM	NRMSE	0.042 58	0.057 26	0.052 32	0.072 23
	Rel. residual	0.016 91	0.018 64	0.019 46	0.048 45
LCA	NRMSE	0.011 56	0.024 12	0.030 97	0.070 07
	Rel. residual	0.017 94	0.013 73	0.019 67	0.015 42
LCM	NRMSE	0.006 18	0.008 69	0.022 33	0.058 67
	Rel. residual	0.004 51	0.004 26	0.011 69	0.033 11

ETNA
Kent State University and
Johann Radon Institute (RICAM)

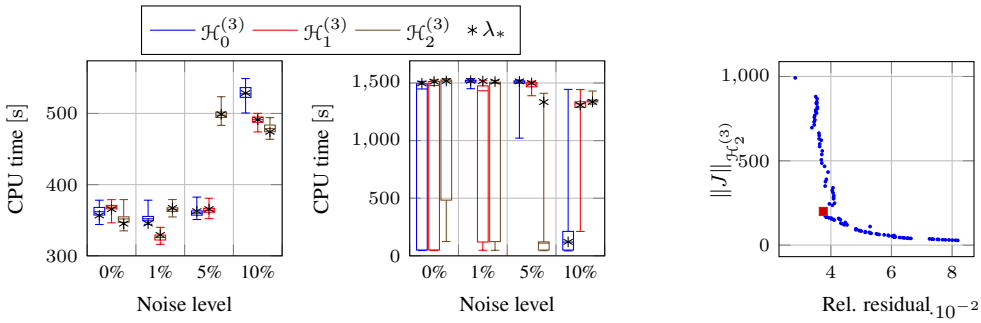

FIG. 8.5. *Boxplots of the required CPU time for the MEG with the RFMP (left) and the ROFMP (middle) problem in seconds and the L-curve in the* 5 % *noise case (right).*

8.3. Comparison with other regularization methods. We compare the results achieved with the ROFMP with two other discretize-then-optimize reconstruction methods. The first method is a simple regularized Ritz method (see [24]), whereas the regularized Tikhonov functional is minimized over an m-dimensional subspace $\mathcal{V}_m$ of $\mathcal{X}$. For instance, in our case, $\mathcal{V}_m$ is spanned by m orthonormal basis functions. In the MEG case, we additionally use the scalar spline method known from MEG data inversion in [19]. For a fair comparison, we modified this method based on reproducing kernels such that the used orthonormal basis functions satisfy the minimum-norm uniqueness constraint. Afterwards, the scalar solution needs to be transferred to the vectorial current; see [35, 37] for this procedure. In the EEG case, the assumptions required for the scalar spline method (see [19]) contradict the minimum-norm solution; see [35]. Hence, we use the vector spline method (see [37]) for the comparison.

TABLE 8.6
NRMSE and relative residual for noise levels achieved by the ROFMP ($s = 2$), the regularized scalar spline method, and the regularized Ritz method ($s = 0$) for the synthetic MEG test case.

	ROFMP		Vector spline		Ritz	
δ_{M}	NRMSE	rel. residual	NRMSE	rel. residual	NRMSE	rel. residual
0.00	$6.477 \cdot 10^{-3}$	$3.596 \cdot 10^{-3}$	$1.272 \cdot 10^{-1}$	$7.482 \cdot 10^{-7}$	$9.260 \cdot 10^{-2}$	$7.899 \cdot 10^{-2}$
0.01	$1.080 \cdot 10^{-2}$	$1.283 \cdot 10^{-2}$	$8.696 \cdot 10^{-2}$	$3.461 \cdot 10^{-3}$	$9.398 \cdot 10^{-2}$	$7.812 \cdot 10^{-2}$
0.05	$4.696 \cdot 10^{-2}$	$3.749 \cdot 10^{-2}$	$1.008 \cdot 10^{-1}$	$4.198 \cdot 10^{-2}$	$8.999 \cdot 10^{-2}$	$8.085 \cdot 10^{-2}$
0.10	$6.043 \cdot 10^{-2}$	$7.638 \cdot 10^{-2}$	$1.092 \cdot 10^{-1}$	$8.172 \cdot 10^{-2}$	$9.993 \cdot 10^{-2}$	$1.220 \cdot 10^{-1}$

The NRMSEs and the relative residuals achieved via the ROFMP, the Ritz method, and the used spline method are given in Table 8.6 (MEG) and Table 8.7 (EEG). The comparison of the reconstruction methods reveals that the NRMSE achieved by the ROFMP is at least one order of magnitude lower than the NRMSE obtained by the other two methods if the noise level is small, that is, $\delta_{\mathrm{E}} \leq 0.01$. While this discrepancy reduces for higher noise levels, the NRMSE achieved by the ROFMP is still smaller than for the other methods. The Ritz method yields comparable results to the scalar spline method and the vector spline method for small noise levels. In contrast, for higher noise levels, the vector spline method is better than the Ritz method with respect to the NRMSE. However, the approximation quality of the ROFMP cannot be reached by either of these two methods.

A visualization of the achieved approximations can also be found in Figure 8.6. Comparing the reconstruction of the Ritz method to the ROFMP reconstruction, the active regions produced by the Ritz method are not properly localized any more, which is outperformed by the results obtained with the ROFMP. In addition, the maximal deviation of the ROFMP solution from

ETNA
Kent State University and
Johann Radon Institute (RICAM)

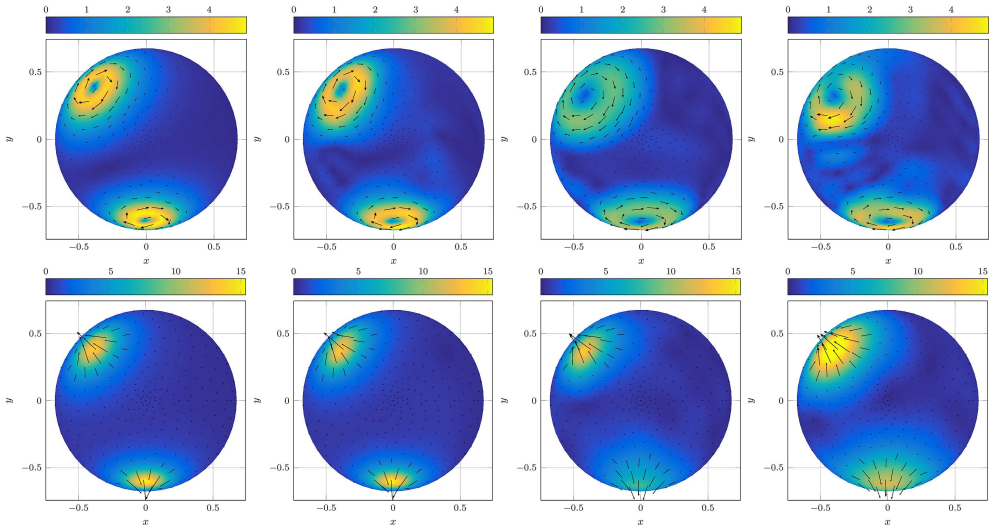

FIG. 8.6. *Approximation of the neuronal current with* $5\,\%$ *noise obtained by the ROFMP (second column), scalar spline method (third column, top), vector spline method (third column, bottom), and the Ritz method (right column) for MEG (top row) and EEG (bottom row) compared to the exact solution (left column).*

the exact solution is only $16.8\,\%$ as high as the deviation of the spline reconstruction. In conclusion, we clearly see that the structure, localization, and order of magnitude of the approximation obtained by the ROFMP are closest to the synthetic test current from Figure 8.2 (left) and Figure 8.4 (left) among all tested methods.

In comparison to the results obtained by the ROFMP from Figure 8.4, we observe that the amplitudes of the reconstruction via the spline method are significantly smaller than those of the ROFMP. Particularly, in the case of the vector spline method, we see that the amplitudes of the spline reconstruction are smaller than those of the ROFMP and the active regions are more spread out, especially for the lower kernel reconstruction. In addition, the maximal deviation of the scalar spline method in the MEG case is significantly damped on both active regions. The maximal deviation achieved via the splines methods is up to 6–7 times higher than achieved with the ROFMP. In addition, the regularization parameter of the spline methods and the Ritz method are chosen by the L-curve method, which yielded good results before; see [37]. However, especially in the case of the scalar spline method, the reconstruction seems over-regularized. Within this method, the regularization parameter is chosen via the L-curve method using (based on the scalar method) the norm of the scalar approximation. It is possible that the detour via a scalar function in the reconstruction causes a larger error here in the reconstruction. In contrast to the spline solution, which is still localized to some extent, the Ritz method yields an even less localized reconstruction. The active regions are much larger than the exact solution, and their amplitudes do not match either. In conclusion, the ROFMP again provides the best reconstruction in terms of the structure, localization, and order of magnitude among all the tested methods.

This accuracy in the reconstruction is at the expense of computation time. In the case of the vector spline method, the CPU time required for the inversion is $90.0000 \pm 3.8299 \times 10^{-4}\,\mathrm{s}$ tested among $10\,000$ inversions. In comparison to a CPU time of $10.5175 \pm 0.2399\,\mathrm{s}$ for the inversion of the EEG data based on the regularized Ritz method throughout our numerical experiments, the vector spline method is slower but still fast when compared to the ROFMP. Note that the CPU time required for the inversion of the MEG data is at least as long as for

ETNA
Kent State University and
Johann Radon Institute (RICAM)

the EEG, since more data is available. The computational time required for the R(O)FMP is plotted in Figure 8.5 and is at least five times as long as the previously mentioned timings, but usually more than tenfold. That does not include the time required for the preprocessing of the dictionary, which is around $6000\,\mathrm{s}$ (wall clock time) for both problems using a parallelized and vectorized code on 12 CPU cores. However, in the case of different data sets (e.g., different patients on the same device), this preprocessing need not be repeated.

TABLE 8.7
NRMSE and relative residual for different noise levels $\delta_{\mathrm{E}} \in \{0, 0.01, 0.05, 0.1\}$ achieved by the ROFMP ($s = 1$), the vectorial spline method, and the Ritz method ($s = 0$) for the synthetic EEG test case.

	ROFMP		Spline		Ritz	
δ_{E}	NRMSE	rel. residual	NRMSE	rel. residual	NRMSE	rel. residual
0.00	$3.487 \cdot 10^{-3}$	$1.513 \cdot 10^{-3}$	$4.160 \cdot 10^{-2}$	$3.457 \cdot 10^{-14}$	$5.766 \cdot 10^{-2}$	$3.780 \cdot 10^{-13}$
0.01	$6.461 \cdot 10^{-3}$	$9.033 \cdot 10^{-3}$	$4.213 \cdot 10^{-2}$	$3.707 \cdot 10^{-14}$	$7.911 \cdot 10^{-2}$	$2.338 \cdot 10^{-3}$
0.05	$1.596 \cdot 10^{-2}$	$1.788 \cdot 10^{-2}$	$4.689 \cdot 10^{-2}$	$3.599 \cdot 10^{-14}$	$1.160 \cdot 10^{-1}$	$2.966 \cdot 10^{-2}$
0.10	$4.700 \cdot 10^{-2}$	$3.866 \cdot 10^{-2}$	$8.623 \cdot 10^{-2}$	$8.328 \cdot 10^{-14}$	$1.170 \cdot 10^{-1}$	$5.527 \cdot 10^{-2}$

Naturally, the question arises whether some methods yield better reconstructions than others. For the regularized Ritz method, the Tikhonov-regularized normal equation is solved over a finite-dimensional subspace spanned by orthonormal polynomials. Thus, local effects, such as certain outliers in the data or locally higher noise levels, have a global impact on the reconstruction. This leads to a general blurriness in the reconstructions for higher noise levels even in the regions where no activity is presumed. In addition, such "global" methods often experience difficulties with irregularly distributed data grids, which are also found in our application; see [43, Sec. 5.3].

The spline-based methods are interpolation methods that satisfy a best approximation property. In the regularized case, which is used for the MEG scalar spline method, the solution of this approach is the unique solution of the Tikhonov-regularized normal equation in a certain Sobolev space. However, via the transformation to the vector-valued case, the structure and quality of the approximation get lost. Thus, the lack of quality of the vector-valued reconstruction presumably originates from the Helmholtz decomposition approach. For this purpose, a direct reconstruction of the vector-valued current should be preferred. This can also be seen in the vector spline method used for the EEG. In Figure 8.6, we see that this method works well if the activity is near the data points, which is the case for the activity in the upper left region of the plot. In contrast, the second active region lies in the middle of a sensor gap, which results in a too flat approximation even in the case of non-noisy data. This is a known disadvantage of interpolation-based approximation methods. Besides, the reconstructions obtained via splines are robust with respect to the noise level. In addition, coarse structures can be reconstructed quickly with the regularized Ritz method, since only a few basis functions are required for this purpose.

The RFMP, as well as its enhancement, the ROFMP, combine the advantages of these methods. Due to the reproducing kernels in the dictionary, which are related to splines, the R(O)FMP is robust with respect to the noise level and can also handle irregularly distributed point grids. Coarse structures can be reconstructed with only a few dictionary elements using the additional orthonormal basis functions. This also results in a sparse solution, especially for lower noise levels, which can be seen in Table 8.3. Using the penalty term of the R(O)FMP, the smoothness of the reconstruction can be controlled, similar to the spline methods. In contrast to the spline methods, the R(O)FMP is not an interpolation method, which results in a better handling of the lack of data. Finally, the R(O)FMP is directly used for the reconstruction of the entire vector-valued neuronal current, which is more stable than a reconstruction of

ETNA
Kent State University and
Johann Radon Institute (RICAM)

scalar-valued parts that are transferred to the vector-valued current afterwards.

8.4. Inversion of real data. After having tested the R(O)FMP extensively within our synthetic test case, we complete this article by reconstructing neuronal currents from real data. A set of real data was gratefully provided to the authors by Dr. O. Hauk (MRC Cognition and Brain Sciences Unit (CBU), Cambridge, UK) and is published in [36]. To generate the data, a human participant wearing an EEG sensor cap was placed into an Elekta Neuromag® [13] MEG device at the MRC CBU. During the measurements, a visual stimulus in the form of a chequered pattern is presented to the participant in the right visual hemi-field (VR, visual right). After a delay, the brain activity increases as well as the magnetic flux density and the electric potential. For the inversion, we choose the measurement corresponding to the point of time with the highest values of the measured quantity. Since the optical nerve fibres associated to the nasal side of the retinas cross each other in the optic chiasm, the brain activity should be maximal at the contralateral visual cortex; see [31]. This means that a visual right stimulus mainly induces brain activity in the left visual cortex. In the case of our three-shell model, the medial longitudinal fissure separating the two brain hemispheres is located in the yz-plane. On the other hand, the position of the face is reflected in the round-shaped gap in the sensor distribution on the front side of the head in the direction of $(0,1,0)^\intercal$.

For the inversion of real data, we use the same setting as within the synthetic test case with $\mathcal{H}_2^{(3)}$-penalty term (MEG). To this end, the ROFMP is started with 500 to 750 different regularization parameter values. After applying all parameter choice methods to the results, we deduce that the L-curve method yields good results by comparing the reconstruction with the measurements. The corresponding L-curve for the real MEG data set is plotted in Figure 8.7 (right). In the EEG case, we test the $\mathcal{H}_1^{(2)}$-norm as well as the $\mathcal{H}_2^{(2)}$-norm. Comparing the reconstructions corresponding to the regularization parameter of the various parameter choice methods, the MGCV yields promising results combined with the $\mathcal{H}_2^{(2)}$-penalty term. However, in the case of the $\mathcal{H}_1^{(2)}$-regularization, all parameter choice methods failed, and we were forced to choose a good regularization parameter manually.

Now, we further investigate the ROFMP runs belonging to the presented parameters for the VR data set. From some applications in the geosciences, it is known that the R(O)FMP tends to first approximate coarse structures with global functions, such as the orthonormal basis functions, and afterwards it reconstructs details with the more localized trial functions; see [15, 17, 49, 57]. In the first 75 iterations, mainly orthonormal basis (ONB) functions and reproducing kernels with a larger width, that is, $h = 0.8$ and sometimes $h = 0.9$, are chosen; see Figure 8.7 (left). Afterwards, smaller reproducing kernels belonging to $h \in \{0.9, 0.95\}$ are chosen primarily. In the end, very fine reproducing kernels belonging to $h = 0.99$ are chosen increasingly. This trend is more pronounced in the real data case than in the synthetic test case, which could be attributed to the richer structure of the real neuronal current. We observe (see Figure 8.7 (middle)) that mainly reproducing kernels located in an outer shell of the cerebrum are chosen, which corresponds to the location of the visual cortex.

In order to get an impression of the quality of the reconstruction, we plotted the absolute value of the measured data separately. In contrast to previous plots, we decided not to show the direction of the neuronal current in order not to overload the plots.

In the case of the MEG recordings, numerical results are depicted in Figure 8.8. On the back of the head, we see that activity in the contralateral left visual cortex is reconstructed, which fits to the recorded data, the experiment, and the theory. However, in the opinion of the authors, some artefacts are reconstructed in the front of the brain. These artefacts are more or less pronounced depending on the chosen regularization parameter. In the case where the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

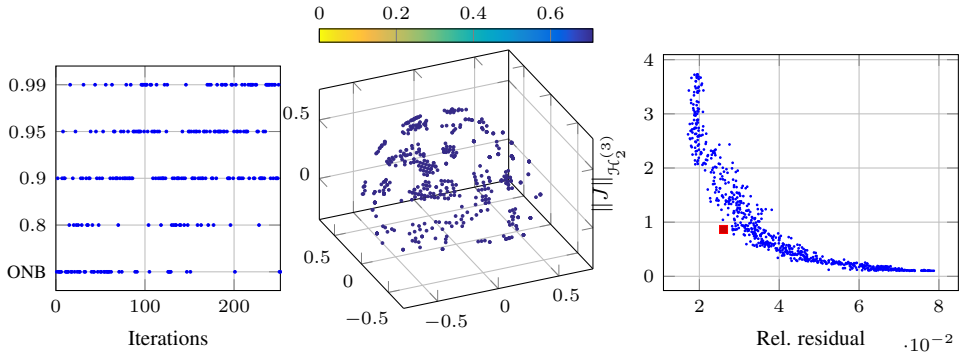

FIG. 8.7. *Chosen dictionary elements (left) depending on reproducing kernel parameter h, kernel centres (middle) by means of the ROFMP (the colour denotes the distance to the origin), and L-curve (right).*

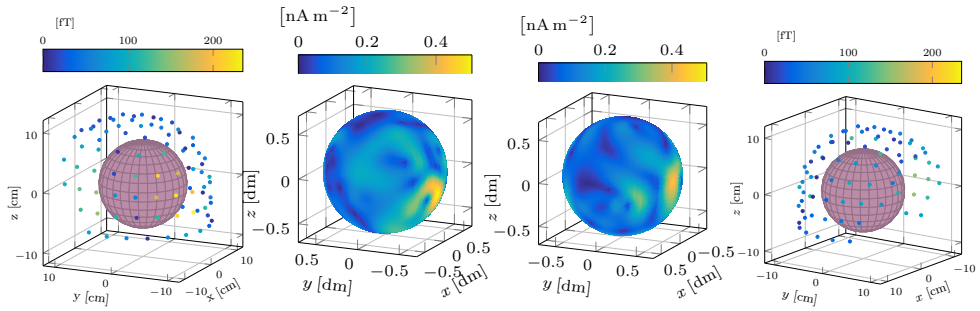

FIG. 8.8. *Real magnetic flux density data (VR) back view (first) and front view (last) and the neuronal current reconstruction via the ROFMP with L-curve from back view (second) and front view (third).*

regularization parameter is chosen according to the LCM, we see that this artefact is located in the data gap in the area of the face.

Inversions of the real electric potential data are visualized in Figure 8.9. The activity of the visual cortex on the left-hand side is clearly recognizable and only small activity at the front is reconstructed. For this inversion, we conclude that the ROFMP is able to reconstruct a reasonable neuronal current, which can be verified, in particular, in Figure 8.9. However, depending on the data set and the chosen penalty term, it is hard to find an appropriate regularization parameter. In contrast, in some tests, activity is reconstructed in areas where no activity is recorded, such as in the case of Figure 8.8. It stands to reason that this activity is an artefact fostered by the lack of data.

We chose the same point in time for the reconstruction of both measurements, which enables us to combine the MEG and EEG reconstructions to a simultaneous inversion. Finally, the sum of the two separate inversions from Figures 8.8 and 8.9 is given in Figure 8.10. The main activity of the brain is located in the visual left cortex, as it should be. In the visualization of the neuronal current on a cutout of the ball, one can additionally observe that most of the activity is located in the outer region of the ball, which corresponds to the structure of the visual cortex. In conclusion, the ROFMP solution for this joint data set yields a plausible reconstruction of the neuronal current. The presumed active regions are clearly reconstructed, and the visualization of the neuronal currents fits to the data. Artefacts, which especially occur in the MEG reconstruction, can be smoothed by the joint inversion.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

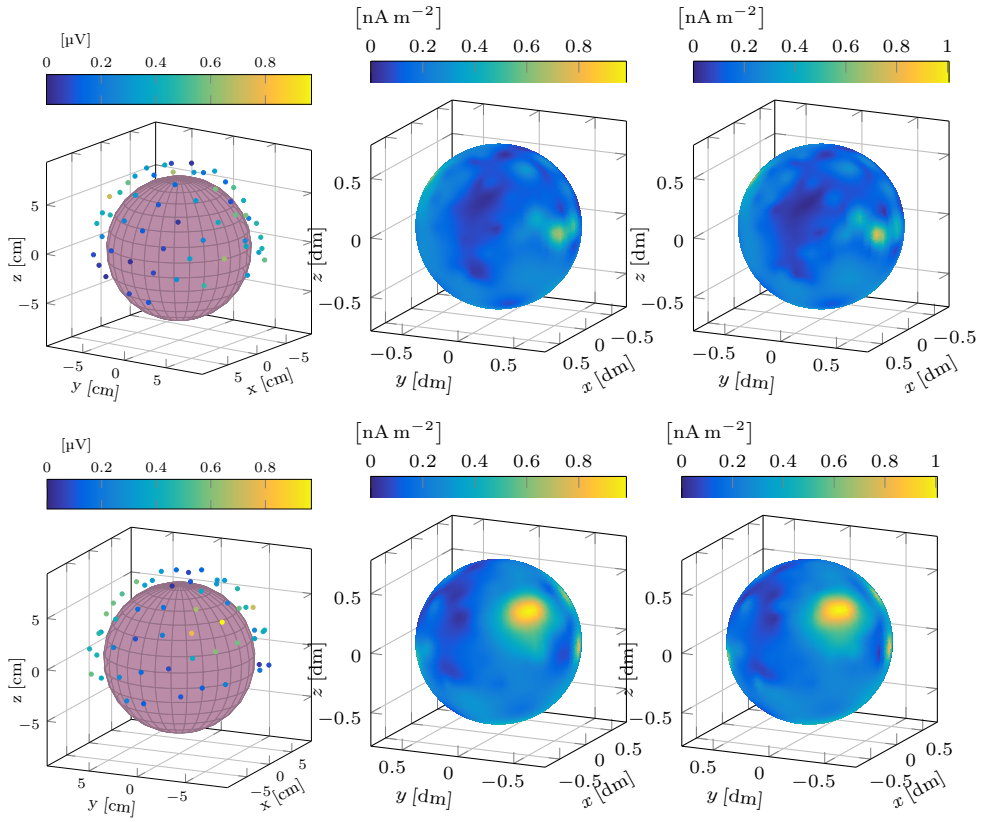

FIG. 8.9. *Real electric potential data (VR) (left column) from front view (top row) and back view (bottom row) and the neuronal current reconstruction via ROFMP for* $\mathcal{H}_2^{(2)}$*-penalty term with MGCV (middle column) and* $\mathcal{H}_1^{(2)}$*-penalty term with manual parameter (third column).*

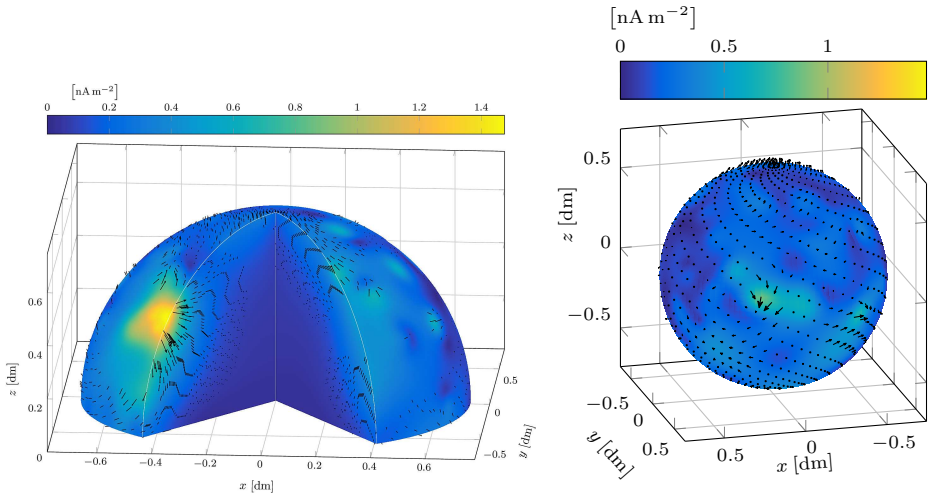

FIG. 8.10. *Neuronal current reconstruction from real magnetic flux density and electric potential data (VR) plotted on a cutout of the cerebrum viewed from the back (left) and on a sphere from the front (right).*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

9. Conclusion and outlook. In this article, we have presented the regularized functional matching pursuit (RFMP) algorithm and its enhancement, the regularized orthogonal functional matching pursuit algorithm (ROFMP). Both iterative algorithms minimize the Tikhonov functional in each step by choosing the next best-fitting function from a so-called dictionary. This type of algorithm can be used for solving severely ill-posed (functional) inverse problems. For this purpose, the penalty term is used for regularization. Under certain conditions on the dictionary, we obtain the convergence of the sequence produced by the RFMP to the best approximate solution of the Tikhonov functional. In addition, we presented how changes in the penalty term (e.g., caused by changing the Sobolev-norm generating pseudo-differential operator) affect the reconstruction.

In order to demonstrate the strength of the R(O)FMP, we apply it to severely ill-posed inverse problems from the context of medical imaging. We reconstruct the neuronal current from non-invasive electric potential (EEG) and magnetic flux density (MEG) measurements. These two inverse problems are ill-posed, since the related null spaces are infinite-dimensional and their singular values decrease towards zero exponentially fast. For the numerics, we bypass the non-uniqueness by considering the additional minimum-norm uniqueness constraint. The instability is opposed by choosing an appropriate penalty term, which is a norm of a spherical Sobolev space, and a regularization parameter by several parameter choice methods.

Due to the non-uniqueness of the EEG and MEG problem, the question arises whether a joint inversion of both data sets yields a more accurate reconstruction than separate inversions. In fact, we are able to prove that no additional information can be gained from the joint inversion, due to the proven complementarity of their null spaces. In addition, due to the possibility to adapt the regularization parameter in the case of the separate inversions in a more data-driven way, the combination of the separate reconstructions should be preferred.

The functionality of the algorithms is demonstrated in synthetic test cases, where the exact solution is known. We have tested non-noisy data and data with additive Gaussian white noise up to 10 %, which is also the supposed noise level for the real data. Within our synthetic test, the ROFMP yields better results than the RFMP with respect to the normalized RMSE on the plotting grid. This difference can be observed, in particular, for non-noisy data and shrinks with increasing noise level. However, both algorithms are able to reconstruct the active regions from the few and irregularly distributed noisy measurements. Within our tests, it reveals that an L^2-penalty term yields poorer results than a Sobolev-norm penalty term, which generated smoother reconstructions and avoided overfitting. The data misfit measured by the relative residual drops below the noise level within a few ROFMP iterations. The reconstructions are stable for increasing noise level. The inversion of the real data yields satisfactory results, taking the physiological expectations and the data situation into account, if a good regularization parameter is found.

We have compared the result achieved via the ROFMP with other optimize-then-discretize regularization methods such as the Ritz method and spline methods. First, the Ritz method solves the regularized Tikhonov-normal equation on a finite-dimensional subspace. Hence, local effects have a global impact on the reconstruction. The produced blurriness is a typical behaviour for such "global" methods on irregularly distributed data grids. Second, the spline method is an interpolation method, which tends to overfit the data. On the other hand, the ROFMP combines the advantages of both methods by using global orthonormal basis functions and localized reproducing kernels in the dictionary. Consequently, coarse trends are reconstructed in the first iterations by global functions, and afterwards details are reconstructed via the aid of localized trial functions.

Within the synthetic test cases, as well as for the real data inversion, it is difficult to find the best (or even a good) regularization parameter. We have tested numerous parameter choice

ETNA
Kent State University and
Johann Radon Institute (RICAM)

methods, where the manual L-curve method and the modified generalized cross validation (MGCV) yielded the best results in most cases. Especially in some of the real data situations, no method was able to find an appropriate regularization parameter. Thus, finding a good regularization parameter for the R(O)FMP is still an open question.

Based on the review of the executed real data inversions, the results of the ROFMP can still be optimized. This enhancement can be performed on different levels. The used dictionary consists of orthonormal basis functions and particular reproducing kernels in our numerical tests. In addition, we could add further types of functions such as appropriate wavelets. Wavelet-based reconstructions and a multi-resolution analysis yield good results in spherical applications; see, e.g., [4, 6, 40, 42]. Besides an enlargement of the dictionary, we could additionally test regularization terms that also take the time dependence of the data into account. Instead of only penalizing the approximation in a certain norm, we could additionally penalize its deviation from that of the previous time step. Second, from the algorithmic side, we can use an enhancements of the R(O)FMP. For instance, in [33, 34], a massively accelerated variant of the algorithm called the regularized weak functional matching pursuit algorithm (RWFMP) is developed, which yielded good results in the considered numerical test cases. Due to the acceleration of the algorithm, more iterations in a shorter period of time are realizable, which is generally desirable. Moreover, in [48, 54, 55], a learned-dictionary-based enhancement of the RFMP is presented, which improves the numerical results by allowing a free choice of the parameter h of the kernels. Lastly, a reason for the discrepancy between the quality of the results obtained in the synthetic test case and in the real data situation may be grounded in insufficient modelling of the brain, since the real structure of the brain is inherently non-spherical. For this purpose, the derived integral equations modelling the inverse MEG and EEG problem need to be adapted, for which further research is required; see [26] for a possible approach.

Acknowledgments. We gratefully acknowledge the support by the German Research Foundation (DFG), projects MI 655/10-1 and MI 655/10-2. We also appreciate the possibility to use the HORUS and the OMNI parallel computing cluster of the University of Siegen, Germany. This paper comprises and extends some results from S.L.'s PhD. thesis; see [35].

Appendix A. Foundations for implementation.

A.1. Proof of Theorem 7.1.

Within this theorem, the forward functionals used in the R(O)FMP are stated. Here, we present the corresponding proof.

Proof of Theorem 7.1. The first two identities concerning the orthonormal basis functions $\tilde{g}^{(i)}_{0,n,j}(\varrho_0;\cdot)$ for $i = 2, 3$ and for all $n \in \mathbb{N}$, $j = 1, \dots, 2n+1$, are based on the results of the SVD of $\mathcal{A}_{\mathrm{M}}$ and $\mathcal{A}_{\mathrm{E}}$ stated in [38, Thms. 6.1 and 6.3], that is,

$$(\mathrm{A.1}) \qquad (\mathcal{A}_{\mathrm{E}}\tilde{g}^{(2)}_{0,n,j}(\varrho_0;\cdot))(y) = \frac{1}{\sqrt{n\varrho_0}}\beta_n^{(L)}\left((n+1)\left(\frac{s}{\varrho_L}\right)^{2n+1} + n\right)\left(\frac{\varrho_0}{s}\right)^{n+1} Y_{n,j}(\eta),$$

$$(\mathcal{A}_{\mathrm{M}}\tilde{g}^{(3)}_{0,n,j}(\varrho_0;\cdot))(y) = -\mu_0\sqrt{\frac{n}{\varrho_0(2n+1)(2n+3)}}\left(\frac{\varrho_0}{s}\right)^{n+2} \tilde{y}^{(1)}_{n,j}(\eta).$$

With the representation of the kernel $k_h^{(3)}$ (see equation (7.2)), the orthonormal basis functions $G^{(i)}_{0,n,j}$, and outer harmonics, we obtain, for all $h \in (0,1)$, $t_n = n$, and $z \in B_{\varrho_0}$ with the sequence $(h_n)_n$ for the symbols of the kernels, the identity

$$(\mathcal{A}_{\mathrm{M}}k_h^{(3)}(\cdot, z))(y)$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$$
\begin{aligned}
&= -\mu_0 \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} h^n \sqrt{\frac{n}{\varrho_0(2n+1)(2n+3)}} \left(\frac{\varrho_0}{s}\right)^{n+2} \tilde{y}_{n,j}^{(1)}(\eta) G_{0,n,j}^{(i)}(\varrho_0; z) \\
&= -\mu_0 \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} h^n \sqrt{\frac{n}{2n+1}} \frac{\varrho_0^n}{s^{n+2}} \left(\frac{v}{\varrho_0}\right)^n P_0^{(0,n+1/2)}\left(2\frac{v^2}{\varrho_0^2} - 1\right) Y_{n,j}(\zeta) \tilde{y}_{n,j}^{(1)}(\eta) \\
&= -\frac{\mu_0}{4\pi} \sum_{n=1}^{\infty} h^n \sqrt{n(2n+1)} \frac{v^n}{s^{n+2}} \tilde{p}_n^{(1)}(\eta, \zeta).
\end{aligned}
$$

We have used the addition theorem equation (see equation (3.2)) in the third step. For the implementation, we need to calculate the vector-valued Legendre function further. We obtain, for all $n \in \mathbb{N}$ and $\eta, \zeta \in S$, and the definition of the vector-valued Legendre polynomials in equation (3.1), the identity

$$
\sqrt{n(2n+1)}\tilde{p}_n^{(1)}(\eta, \zeta) = \sqrt{n(n+1)}\eta P_n(\eta \cdot \zeta) - \sqrt{\frac{n}{n+1}}(\zeta - (\eta \cdot \zeta)\eta) P_n'(\eta \cdot \zeta).
$$

We use similar techniques and derive, for the kernel $k_h^{(2)}$, the identity

$$
\begin{aligned}
&(\mathcal{A}_{\mathrm{E}} k_h^{(2)}(\cdot, z))(y) \\
&\quad = \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} h^n \frac{1}{\sqrt{n}\varrho_0} \beta_n^{(L)} \left((n+1)\left(\frac{s}{\varrho_L}\right)^{2n+1} + n\right) \left(\frac{\varrho_0}{s}\right)^{n+1} Y_{n,j}(\eta) G_{0,n,j}^{(i)}(\varrho_0; z) \\
&\quad = \frac{1}{4\pi} \sum_{n=1}^{\infty} h^n \sqrt{\frac{(2n+1)^3}{n}} \beta_n^{(L)} \left((n+1)\left(\frac{s}{\varrho_L}\right)^{2n+1} + n\right) \frac{v^{n-1}}{s^{n+1}} P_n(\eta \cdot \zeta),
\end{aligned}
$$

for all $h \in (0, 1)$ and $z \in B_{\varrho_0}$. Note that in this case the sequence $\{t_n\}_{n \in \mathbb{N}}$ occurring in the functions $G_{0,n,j}^{(i)}(\varrho_0; \cdot)$ is given by $t_n := n - 1$ for all $n \in \mathbb{N}$. ∎

A.2. Proof of Theorem 7.2. Here, the inner products of the dictionary elements with respect to the used Sobolev norms are stated. They can be calculated as follows.

Proof of Theorem 7.2. For the orthonormal basis functions, we get, for all $n, \bar{n} \in \mathbb{N}$, $j = 1, \ldots, 2n+1$, and $\bar{j} = 1, \ldots, 2\bar{n}+1$, via Parseval's identity and equation (7.1), the relation

$$
\begin{aligned}
&\left\langle \tilde{g}_{0,n,j}^{(i)}(\varrho_0; \cdot), \tilde{g}_{0,\bar{n},\bar{j}}^{(i)}(\varrho_0; \cdot) \right\rangle_{\mathcal{H}_\bullet^s} \\
&\quad = \sum_{\substack{(m',n') \in \mathbb{N}_0 \times \mathbb{N} \\ a_{m',n'}^{(i)} \neq 0}} \sum_{j'=1}^{2n'+1} (a_{m',n'}^{(i)})^2 \left\langle \tilde{g}_{0,n,j}^{(i)}(\varrho_0; \cdot), \tilde{g}_{m',n',j'}^{(i)}(\varrho_0; \cdot) \right\rangle_{\mathrm{L}^2(B_{\varrho_0})} \\
&\qquad \times \left\langle \tilde{g}_{0,\bar{n},\bar{j}}^{(i)}(\varrho_0; \cdot), \tilde{g}_{m',n',j'}^{(i)}(\varrho_0; \cdot) \right\rangle_{\mathrm{L}^2(B_{\varrho_0})} \\
&\quad = (a_n^{(i)})^2 \left\langle \tilde{g}_{0,\bar{n},\bar{j}}^{(i)}(\varrho_0; \cdot), \tilde{g}_{0,n,j}^{(i)}(\varrho_0; \cdot) \right\rangle_{\mathrm{L}^2(B_{\varrho_0})} = (a_n^{(i)})^2 \delta_{n,\bar{n}} \delta_{j,\bar{j}}.
\end{aligned}
$$

In order to calculate the $\mathcal{H}_\bullet^s(a, B_{\varrho_0})$-inner product of reproducing kernels localized in $z_l \in B_{\varrho_0}$ for $l = 1, \ldots, L_\bullet$ with other dictionary elements, we compute an auxiliary inner product. For all $n \in \mathbb{N}$, $j = 1, \ldots, 2n+1$, and $h \in (0, 1)$, we have

$$
\left\langle k_h^{(i)}(\cdot, z_l), \tilde{g}_{0,n,j}^{(i)}(\varrho_0; \cdot) \right\rangle_{\mathrm{L}^2(B_{\varrho_0})}
$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$$= \left\langle \sum_{n'=1}^{\infty} \sum_{j'=1}^{2n'+1} h^{n'} G^{(i)}_{0,n',j'}(\varrho_0; z_l) \tilde{g}^{(i)}_{0,n',j'}(\varrho_0; \cdot), \tilde{g}^{(i)}_{0,n,j}(\varrho_0; \cdot) \right\rangle_{\mathrm{L}^2(B_{\varrho_0})}$$
$$= h^n G^{(i)}_{0,n,j}(\varrho_0; z_l).$$

Eventually, we obtain via Parseval's identity and the definition of the orthonormal basis functions, for all $z_l \in B_{\varrho_0}$, $l = 1, \dots, L_\bullet$, $h \in (0,1)$, $n \in \mathbb{N}$, and $j = 1, \dots, 2n+1$, the relation

$$\begin{aligned}
&\left\langle k^{(i)}_h(\cdot, z_l), \tilde{g}^{(i)}_{0,n,j}(\varrho_0; \cdot) \right\rangle_{\mathcal{H}^s_\bullet} \\
&\quad = \sum_{\substack{(m',n') \in \mathbb{N}_0 \times \mathbb{N} \\ a^{(i)}_{m',n'} \neq 0}} \sum_{j'=1}^{2n'+1} (a^{(i)}_{m',n'})^2 \left\langle k^{(i)}_h(\cdot, z_l), \tilde{g}^{(i)}_{m',n',j'}(\varrho_0; \cdot) \right\rangle_{\mathrm{L}^2(B_{\varrho_0})} \delta_{m',0} \delta_{n',n} \delta_{j',j} \\
&\quad = (a^{(i)}_n)^2 \left\langle k^{(i)}_h(\cdot, z_l), \tilde{g}^{(i)}_{0,n,j}(\varrho_0; \cdot) \right\rangle_{\mathrm{L}^2(B_{\varrho_0})} = (a^{(i)}_n)^2 h^n G^{(i)}_{0,n,j}(\varrho_0; z_l) \\
&\quad = \sqrt{\frac{2t^{(i)}_n + 3}{\varrho_0^3}} (a^{(i)}_n)^2 h^n \left(\frac{v_l}{\varrho_0} \right)^{t^{(i)}_n} Y_{n,j}(\zeta_l).
\end{aligned}$$

Finally, for two kernel functions with centres $z_l, z_k \in B_{\varrho_0}$ where $l, k \in \{1, \dots, L_\bullet\}$ and parameters $h, h' \in (0,1)$, we get, with the precise representation of the orthonormal basis, the result

$$\begin{aligned}
&\left\langle k^{(i)}_h(\cdot, z_l), k^{(i)}_{h'}(\cdot, z_k) \right\rangle_{\mathcal{H}^s_\bullet} \\
&\quad = \sum_{\substack{(m,n) \in \mathbb{N}_0 \times \mathbb{N} \\ a^{(i)}_{m,n} \neq 0}} \sum_{j=1}^{2n+1} (a^{(i)}_{m,n})^2 \left\langle k^{(i)}_h(\cdot, z_l), \tilde{g}^{(i)}_{m,n,j}(\varrho_0; \cdot) \right\rangle_{\mathrm{L}^2(B_{\varrho_0})} \\
&\qquad \times \left\langle k^{(i)}_{h'}(\cdot, z_k), \tilde{g}^{(i)}_{m,n,j}(\varrho_0; \cdot) \right\rangle_{\mathrm{L}^2(B_{\varrho_0})} \\
&\quad = \sum_{\substack{n \in \mathbb{N} \\ a^{(i)}_n \neq 0}} \sum_{j=1}^{2n+1} (a^{(i)}_n)^2 (hh')^n G^{(i)}_{0,n,j}(\varrho_0; z_l) G^{(i)}_{0,n,j}(\varrho_0; z_k) \\
&\quad = \frac{1}{4\pi \varrho_0^3} \sum_{\substack{n \in \mathbb{N} \\ a^{(i)}_n \neq 0}} (a^{(i)}_n)^2 (hh')^n (2n+1)(2t^{(i)}_n + 3) \left(\frac{v_l v_k}{\varrho_0^2} \right)^{t^{(i)}_n} P_n(\zeta_l \cdot \zeta_k). \qquad \square
\end{aligned}$$

Appendix B. Foundation for synthetic test cases.

B.1. Proof of Theorem 7.4.

Theorem 7.4. Here, we need to prove that the test current is entirely contained in the orthogonal complement of the operator null space.

Proof of Theorem 7.4. Since the Legendre polynomials are bounded (see [51, Ch. II.7]) and $hrv/\varrho_0^2 < 1$, the series occurring in the synthetic currents can be estimated by convergent power series. Thus, the series and its derivatives converge uniformly, and we are able to interchange the gradient and the L^* operator with the respective series.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Now, we start with the MEG case. We interchange the differential operator with the series, use the definition of the Morse–Feshbach vector Legendre polynomials and equation (3.2). Hence, for all $x \in B_{\varrho_0}$, we get the identity

$$\begin{aligned} I_{\mathrm{M}}(x;h,z) &= \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \left(\frac{hrv}{\varrho_0^2}\right)^n L_\xi^* P_n(\xi\cdot\zeta) = \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \sqrt{\tilde{\mu}_n^{(3)}} \left(\frac{hrv}{\varrho_0^2}\right)^n \tilde{p}_n^{(3)}(\xi;\zeta) \\ &= \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} \left(\frac{hrv}{\varrho_0^2}\right)^n \sqrt{\tilde{\mu}_n^{(3)}}\, \tilde{y}_{n,j}^{(3)}(\xi) Y_{n,j}(\zeta) \\ &= \varrho_0^3 \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} h^n \frac{\sqrt{n(n+1)}}{2n+3} \tilde{g}_{0,n,j}^{(3)}(\varrho_0;x) G_{0,n,j}^{(3)}(\varrho_0;z). \end{aligned}$$

In the last step, we used the definition of the scalar-valued and vector-valued basis functions with $m = 0$ and $t_n^{(3)} = n$.

In the EEG case, we analogously obtain with equation (3.1), but for $m = 0$, $t_n^{(2)} = n-1$, and $t_n^{(3)} = n$, for all $x \in B_{\varrho_0}$, the representation

$$\begin{aligned} I_{\mathrm{E}}(x;h,z) &= \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \nabla_x \left(\left(\frac{hrv}{\varrho_0^2}\right)^n P_n(\xi\cdot\zeta) \right) \\ &= \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \sqrt{\tilde{\mu}_n^{(2)}} \left(\frac{hv}{\varrho_0^2}\right)^n r^{n-1} \tilde{p}_n^{(2)}(\xi;\zeta) \\ &= \varrho_0^2 \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} h^n \sqrt{\frac{(2n+1)n}{\varrho_0^6}} \frac{v^n}{\varrho_0^n} \frac{r^{n-1}}{\varrho_0^{n-1}} \tilde{y}_{n,j}^{(2)}(\xi) Y_{n,j}(\zeta) \\ &= \varrho_0^2 \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} h^n \sqrt{\frac{n}{2n+3}} \tilde{g}_{0,n,j}^{(2)}(\varrho_0;x) G_{0,n,j}^{(3)}(\varrho_0;z). \end{aligned}$$

Eventually, Theorem 5.1 provides us with the desired result. ∎

B.2. Proof of Lemma 7.5. In the following proof, we verify the closed representation of the synthetic currents.

Proof of Lemma 7.5. We start with the definition of the synthetic current from Definition 7.3, use the closed representation of the Abel–Poisson kernel from equation (7.3), and use the fact that $L^*1 = 0$. Thus, with $q := hrv/\varrho_0^2$, we get, for all $h \in [0,1)$ and $z \in B_{\varrho_0}$, the identity

$$I_{\mathrm{M}}(x;h,z) = \frac{1}{4\pi} L_\xi^* \left(\sum_{n=0}^{\infty} (2n+1) q^n P_n(\xi\cdot\zeta) - 1 \right) = \frac{1}{4\pi} L_\xi^* \frac{1-q^2}{(1+q^2-2q(\xi\cdot\zeta))^{3/2}}.$$

The variable q is independent of the angular part of x. With $L_\xi^* F(\xi\cdot\eta) = F'(\xi\cdot\eta)(\xi\wedge\eta)$ (see [22, eq. (2.152)]) and the quotient rule, we obtain the desired representation.

Now, we use the same considerations as in the MEG case and immediately get for the EEG test current

$$I_{\mathrm{E}}(x;h,z) = \frac{1}{4\pi} \nabla_x \frac{1-q(r)^2}{(1+q(r)^2-2q(r)(\xi\cdot\zeta))^{3/2}}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

The gradient can be split into a radial and an angular derivative. Then we obtain, with $\nabla_\xi^* F(\xi\cdot\eta) = F'(\xi\cdot\eta)(\eta - (\xi\wedge\eta)\xi)$ (see [22, eq. (2.151)]) and the quotient rule, the representation

$$\begin{aligned}4\pi I_{\mathrm{E}}(x;h,z) &= \left(\xi\frac{\partial}{\partial r} + \frac{1}{r}\nabla_\xi^*\right)\frac{1-q(r)^2}{\left(1+q(r)^2-2q(r)(\xi\cdot\zeta)\right)^{3/2}}\\ &= \xi\frac{\partial}{\partial r}\frac{1-q(r)^2}{\left(1+q(r)^2-2q(r)(\xi\cdot\zeta)\right)^{3/2}}\\ &\quad + \frac{3}{r}\frac{(1-q(r)^2)q(r)}{\left(1+q(r)^2-2q(r)(\xi\cdot\zeta)\right)^{5/2}}\left(\zeta-(\xi\cdot\zeta)\xi\right).\end{aligned}$$

In addition, again with the quotient rule, $q'(r) = q(r)/r$, and the abbreviation for p, we get

$$\begin{aligned}&\frac{\partial}{\partial r}\frac{1-q(r)^2}{\left(1+q(r)^2-2q(r)p\right)^{3/2}}\\ &\quad= \frac{q(r)}{r}\frac{-2q(r)(1+q(r)^2-2q(r)p) - 3(1-q(r)^2)\,(q(r)-p)}{\left(1+q(r)^2-2q(r)p\right)^{5/2}}\\ &\quad= \frac{q(r)}{r}\frac{-5q(r)+q(r)^3+q(r)^2p+3p}{\left(1+q(r)^2-2q(r)p\right)^{5/2}}.\end{aligned}$$

Inserting this into the formula for the synthetic current, we get, with $q := q(r)$, the stated result, that is,

$$\begin{aligned}4\pi I_{\mathrm{E}}(x;h,z) &= \frac{q^2(-5+q^2+qp)+3qp}{r\left(1+q^2-2qp\right)^{5/2}}\xi + \frac{3(1-q^2)q}{r\left(1+q^2-2qp\right)^{5/2}}(\zeta-p\xi)\\ &= \frac{q^2(-5+q^2+qp)+3qp-3qp+3q^3p}{r\left(1+q^2-2qp\right)^{5/2}}\xi + \frac{3(1-q^2)q}{r\left(1+q^2-2qp\right)^{5/2}}\zeta\\ &= \frac{q^2(-5+q^2+4qp)\xi+3(1-q^2)q\zeta}{r\left(1+q^2-2qp\right)^{5/2}}.\end{aligned}$$

□

B.3. Proof of Theorem 7.6. Now, we verify the representation of the forward functionals applied to the synthetic test current.

Proof of Theorem 7.6. Via equation (A.1), we get, with the Fourier expression of $J_\bullet$, the series

$$\begin{aligned}(\mathcal{A}_{\mathrm{M}}J_{\mathrm{M}})(y) &= -\mu_0\sum_{n=1}^{\infty}\sum_{j=1}^{2n+1} J_{\mathrm{M}}{}^\wedge(3,0,n,j)\sqrt{\frac{n}{\varrho_0(2n+1)(2n+3)}}\left(\frac{\varrho_0}{s}\right)^{n+2}\tilde{y}_{n,j}^{(1)}(\eta),\\ (\mathcal{A}_{\mathrm{E}}J_{\mathrm{E}})(y) &= \sum_{n=1}^{\infty}\sum_{j=1}^{2n+1} J_{\mathrm{E}}{}^\wedge(2,0,n,j)\frac{1}{\sqrt{n\varrho_0}}\beta_n^{(L)}\left((n+1)\left(\frac{s}{\varrho_L}\right)^{2n+1}+n\right)\\ &\quad\times\left(\frac{\varrho_0}{s}\right)^{n+1}Y_{n,j}(\eta).\end{aligned}$$

Only the calculation of the Fourier coefficients remains to be done. Since J_{M} and J_{E} are linear combinations of $I_{\mathrm{M}}(\cdot,h,z)$ or $I_{\mathrm{E}}(\cdot,h,z)$, respectively, with some parameters $h\in[0,1)$ and $z\in B_{\varrho_0}$, we calculate the Fourier coefficients for these functions and then use the linearity of $\mathcal{A}_{\mathrm{M}}$ and $\mathcal{A}_{\mathrm{E}}$ in order to obtain the result for J_{M} and J_{E}, respectively. By means of the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

representation of the synthetic current from the proof of Theorem 7.4 for $h \in [0,1)$ and $z \in B_{\varrho_0}$, we obtain, for all $n \in \mathbb{N}$ and $j = 1, \ldots, 2n+1$, the identities

$$(I_{\mathrm{M}}(\cdot;h,z))^{\wedge}(3,0,n,j) = \varrho_0^3 h^n \frac{\sqrt{n(n+1)}}{2n+3} G^{(3)}_{0,n,j}(\varrho_0;z), \quad t_n^{(3)} = n,$$
$$(I_{\mathrm{E}}(\cdot;h,z))^{\wedge}(2,0,n,j) = \varrho_0^2 h^n \sqrt{\frac{n}{2n+3}} G^{(3)}_{0,n,j}(\varrho_0;z), \qquad t_n^{(2)} = n-1.$$

In the MEG case, we then obtain, with $h \in [0,1)$ and $z \in B_{\varrho_0}$, that

$$(\mathcal{A}_{\mathrm{M}}(I_{\mathrm{M}}(\cdot;h,z)))(y) = -\mu_0 \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} \sqrt{\frac{n^2(n+1)\varrho_0}{(2n+1)(2n+3)^3}} \left(\frac{\varrho_0}{s}\right)^{n+2} \varrho_0^2 h^n G^{(3)}_{0,n,j}(\varrho_0;z) \tilde{y}^{(1)}_{n,j}(\eta).$$

Using the representation of the orthonormal basis, equation (3.3), and the addition theorem in equation (3.2), we obtain, for $h \in [0,1)$, $z \in B_{\varrho_0}$, and all $y \in \overline{B^{\mathrm{ext}}_{\varrho_L}}$, the relation

$$\begin{aligned}(\mathcal{A}_{\mathrm{M}}(I_{\mathrm{M}}(\cdot;h,z)))(y) &= -\mu_0 \varrho_0^3 \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} \sqrt{\frac{n^2(n+1)}{(2n+1)(2n+3)^2}} \frac{(hv)^n}{s^{n+2}} Y_{n,j}(\zeta) \tilde{y}^{(1)}_{n,j}(\eta) \\ &= -\frac{\mu_0}{4\pi} \varrho_0^3 \sum_{n=1}^{\infty} \sqrt{\frac{n^2(n+1)(2n+1)}{(2n+3)^2}} \frac{(hv)^n}{s^{n+2}} \tilde{p}^{(1)}_n(\eta;\zeta).\end{aligned}$$

In the EEG case, we obtain, with equation (3.2), for all $y \in S_{[\varrho_{L-1},\varrho_L]}$, the result

$$(\mathcal{A}_{\mathrm{E}} I_{\mathrm{E}}(\cdot;h,z))(y) = \frac{\varrho_0}{4\pi} \sum_{n=1}^{\infty} \beta_n^{(L)} \frac{h^n v^n}{s^{n+1}} (2n+1) \left((n+1)\left(\frac{s}{\varrho_L}\right)^{2n+1} + n\right) P_n(\zeta \cdot \eta). \qquad \square$$

B.4. Proof of Lemma 7.7.

Proof of Lemma 7.7. From Theorem 7.6 and the estimate $|\tilde{p}_n^{(i)}| \leq \sqrt{\mu_n^{(i)}}$, we get for the truncation error

$$\frac{\mu_0}{4\pi} \varrho_0^3 \left| \sum_{n=N+1}^{\infty} \sqrt{\frac{n^2(n+1)(2n+1)}{(2n+3)^2}} \frac{(hv)^n}{s^{n+2}} \tilde{p}^{(1)}_n(\eta;\zeta) \right| \leq \frac{\mu_0}{4\pi} \frac{\varrho_0^3}{\varrho_L^2} \sum_{n=N+1}^{\infty} n(n+1) \left(\frac{h\varrho_0}{\varrho_L}\right)^n.$$

We define $q \coloneqq h\varrho_0/\varrho_L \in [0,1)$ and obtain, with the closed representation of the power series and some lengthy calculations, the estimate

$$\begin{aligned}\sum_{n=N+1}^{\infty} n(n+1) \left(\frac{h\varrho_0}{\varrho_L}\right)^n &= \frac{q^{N+1}(-q^2N^2 + 2qN^2 - N^2 - q^2N + 4qN - 3N - 2)}{(q-1)^3} \\ &= \frac{q^{N+1}\left(-(q-1)N((q-1)N + q - 3) - 2\right)}{(q-1)^3} \\ &= \frac{q^{N+1}\left((1-q)N((1-q)N - q + 3) + 2\right)}{(1-q)^3} \\ &\leq \frac{q^{N+1}\left((1-q)N(N+3) + 2\right)}{(1-q)^3} \leq \frac{q^{N+1}(N+2)(N+1)}{(1-q)^3},\end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where $N \in \mathbb{N}$. Vice versa, from Theorem 7.6 with

$$\limsup_{k\to\infty} |k\beta_k^{(L)}| = \frac{1}{2\sigma_L C^{(L)}}$$

(see [38, Lem. 4.2]), the abbreviation $q \coloneqq h\varrho_0/\varrho_{L-1} \in [0,1)$, and the inequality $(2n+1)^2 \leq 9n^2$ for all $n \in \mathbb{N}$, we similarly obtain, for all $N \in \mathbb{N}$, the estimate

$$\begin{aligned}
&\frac{\varrho_0}{4\pi}\left|\sum_{n=N+1}^{\infty} \beta_n^{(L)} \frac{h^n v^n}{s^{n+1}}(2n+1)\left((n+1)\left(\frac{s}{\varrho_L}\right)^{2n+1}+n\right)P_n(\zeta\cdot\eta)\right| \\
&\quad\leq \frac{\varrho_0}{4\pi}\sum_{n=N+1}^{\infty} |\beta_n^{(L)}| \frac{h^n\varrho_0^n}{\varrho_{L-1}^{n+1}}(2n+1)^2 \lesssim \frac{9\varrho_0}{8\pi\sigma_L\varrho_{L-1}C^{(L)}}\sum_{n=N+1}^{\infty} nq^n \\
&\quad= \frac{9\varrho_0}{8\pi\sigma_L\varrho_{L-1}C^{(L)}}\frac{q^{N+1}\left((1-q)N+1\right)}{(1-q)^2} \leq \frac{9\varrho_0}{8\pi\sigma_L\varrho_{L-1}C^{(L)}}\frac{q^{N+1}(N+1)}{(1-q)^2},
\end{aligned}$$

where we have written "$\lesssim$" here because of the estimate with the limes superior. □

REFERENCES

[1] F. BAUER, M. GUTTING, AND M. A. LUKAS, *Evaluation of parameter choice methods for regularization of ill-posed problems in geomathematics*, in Handbook of Geomathematics, W. Freeden, M. Z. Nashed, and T. Sonar, eds., Springer, Berlin, 2015, pp. 1713–1774.

[2] F. BAUER AND M. A. LUKAS, *Comparing parameter choice methods for regularization of ill-posed problems*, Math. Comput. Simulation, 81 (2011), pp. 1795–1841.

[3] H. BAUER, *Measure and Integration Theory*, de Gruyter, Berlin, 2001.

[4] M. BAYER, *Geomagnetic Field Modelling from Satellite Data by First and Second Generation Vector Wavelets*, PhD. Thesis, University of Kaiserslautern, Department of Mathematics, Geomathematics Group, Shaker-Verlag, Aachen, 2000.

[5] P. BERKEL, D. FISCHER, AND V. MICHEL, *Spline multiresolution and numerical results for joint gravitation and normal-mode inversion with an outlook on sparse regularisation*, GEM Int. J. Geomath., 1 (2011), pp. 167–204.

[6] S. BETH, *Multiscale Approximation by Vector Radial Basis Functions on the Sphere*, PhD. Thesis, University of Kaiserslautern, Department of Mathematics, Geomathematics Group, Shaker-Verlag, Aachen, 2000.

[7] J. CHEN, S. PEREVERZYEV, JR., AND Y. XU, *Aggregation of regularized solutions from multiple observation models*, Inverse Problems, 31 (2015), Art. 075005, 23 pages.

[8] C. W. CLENSHAW, *A note on the summation of Chebyshev series*, Math. Tables Aids Comput., 9 (1955), pp. 118–120.

[9] M. CLERC, J. LEBLOND, J.-P. MARMORAT, AND C. PAPAGEORGAKIS, *Uniqueness result for an inverse conductivity recovery problem with application to EEG*, Rend. Istit. Mat. Univ. Trieste, 48 (2016), pp. 385–406.

[10] J. C. DE MUNCK AND M. J. PETERS, *A fast method to compute the potential in the multisphere model (EEG application)*, IEEE Trans. Biomed. Eng., 40 (1993), pp. 1166–1174.

[11] J. R. DRISCOLL AND D. M. HEALY, JR., *Computing Fourier transforms and convolutions on the 2-sphere*, Adv. in Appl. Math., 15 (1994), pp. 202–250.

[12] A. R. EDMONDS, *Angular Momentum in Quantum Mechanics*, Princeton University Press, Princeton, 1957.

[13] ELEKTA NEUROMAG, *Elekta Neuromag System Hardware Technical Manual*, September 2005.

[14] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

[15] D. FISCHER, *Sparse Regularization of a Joint Inversion of Gravitational Data and Normal Mode Anomalies*, PhD. Thesis, University of Siegen, Department of Mathematics, Geomathematics Group, Dr. Hut, Munich, 2011.

[16] D. FISCHER AND V. MICHEL, *Sparse regularization of inverse gravimetry – case study: spatial and temporal mass variations in South America*, Inverse Problems, 28 (2012), Art. 065012, 34 pages.

[17] ———, *Automatic best-basis selection for geophysical tomographic inverse problems*, Geophys. J. Internat., 193 (2013), pp. 1291–1299.

[18] ———, *Inverting GRACE gravity data for local climate effects*, J. Geo. Sci., 3 (2013), pp. 151–162.

[19] A. S. FOKAS, O. HAUK, AND V. MICHEL, *Electro-magneto-encephalography for the three-shell model: numerical implementation via splines for distributed current in spherical geometry*, Inverse Problems, 28 (2012), Art. 035009, 28 pages.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[20] W. FREEDEN, V. MICHEL, AND F. J. SIMONS, *Spherical harmonics based special function systems and constructive approximation methods*, in Handbook of Mathematical Geodesy, W. Freeden and M. Z. Nashed, eds., Geosyst. Math., Birkhäuser/Springer, Cham, 2018, pp. 753–819.

[21] W. FREEDEN AND M. SCHREINER, *Spherical Functions of Mathematical Geosciences. A Scalar, Vectorial and Tensorial Setup*, Springer, Berlin, 2009.

[22] ———, *Special functions in mathematical geosciences: an attempt at a categorization*, in Handbook of Geomathematics, W. Freeden, M. Z. Nashed, and T. Sonar, eds., Springer, Berlin, 2010, pp. 925–948.

[23] C. GERHARDS, S. PEREVERZYEV, JR., AND P. TKACHENKO, *A parameter choice strategy for the inversion of multiple observations*, Adv. Comput. Math., 43 (2017), pp. 101–112.

[24] C. W. GROETSCH, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, 1984.

[25] M. GUTTING, B. KRETZ, V. MICHEL, AND R. TELSCHOW, *Study on parameter choice methods for the RFMP with respect to downward continuation*, Front. Appl. Math. Stat., 3 (2017), Art. 10, 17 pages.

[26] P. HASHEMZADEH, A. S. FOKAS, AND C. B. SCHÖNLIEB, *A hybrid analytical–numerical algorithm for determining the neuronal current via electroencephalography*, J. Royal Soc. Interface, 17 (2020), Art. 20190831, 14 pages.

[27] B. HE, ed., *Modeling and Imaging of Bioelectrical Activity: Principles and Applications*, Kluwer, New York, 2005.

[28] N. J. HIGHAM, *The accuracy of floating point summation*, SIAM J. Sci. Comput., 14 (1993), pp. 783–799.

[29] A. ISHTIAQ, *Grid Points and Generalized Discrepancies on the d-Dimensional Ball*, PhD. Thesis, University of Siegen, Department of Mathematics, Geomathematics Group, 2018.

[30] A. ISHTIAQ AND V. MICHEL, *Pseudo-differential operators, cubature and equidistribution on the 3D ball: an approach based on orthonormal basis systems*, Numer. Funct. Anal. Optim., 38 (2017), pp. 891–910.

[31] E. R. KANDEL, J. H. SCHWARTZ, T. M. JESSELL, S. A. SIEGELBAUM, AND A. J. HUDSPETH, *Principles of Neural Science*, 5th ed., McGraw-Hill, New York, 2013.

[32] K.-R. KOCH AND J. KUSCHE, *Regularization of geopotential determination from satellite data by variance components*, J. Geod., 76 (2002), pp. 259–268.

[33] M. KONTAK, *Novel Algorithms of Greedy-type for Probability Density Estimation as well as Linear and Nonlinear Inverse Problems*, PhD. Thesis, University of Siegen, Department of Mathematics, Geomathematics Group, 2018.

[34] M. KONTAK AND V. MICHEL, *The regularized weak functional matching pursuit for linear inverse problems*, J. Inverse Ill-Posed Probl., 27 (2019), pp. 317–340.

[35] S. LEWEKE, *The Inverse Magneto-Electroencephalography Problem for the Spherical Multiple-Shell Model – Theoretical Investigations and Numerical Aspects*, PhD. Thesis, University of Siegen, Department of Mathematics, Geomathematics Group, 2018.

[36] ———, *rkhs-splines*, Software, 2021. Github repository https://github.com/SarahLeweke/rkhs-splines.

[37] S. LEWEKE, O. HAUK, AND V. MICHEL, *Vector-valued spline method for the spherical multiple-shell electro-magnetoencephalography problem*, Inverse Problems, 38 (2022), Art. 085001, 46 pages.

[38] S. LEWEKE, V. MICHEL, AND A. S. FOKAS, *Electro-magnetoencephalography for a spherical multiple-shell model: novel integral operators with singular-value decompositions*, Inverse Problems, 36 (2020), Art. 035003, 31 pages.

[39] A. K. LOUIS, *Inverse und schlecht gestellte Probleme*, Teubner, Stuttgart, 1989.

[40] T. MAIER, *Multiscale Geomagnetic Field Modelling from Satellite Data: Theoretical Aspects and Numerical Application*, PhD. Thesis, University of Kaiserslautern, Department of Mathematics, Geomathematics Group, 2003.

[41] S. G. MALLAT AND Z. ZHANG, *Matching pursuits with time–frequency dictionaries*, IEEE Trans. Signal. Process., 41 (1993), pp. 3397–3415.

[42] C. MAYER, *Wavelet decomposition of spherical vector fields with respect to sources*, J. Fourier Anal. Appl., 12 (2006), pp. 345–369.

[43] V. MICHEL, *Lectures on Constructive Approximation. Fourier, Spline, and Wavelet Methods on the Real Line, the Sphere, and the Ball*, Birkhäuser, New York, 2013.

[44] ———, *RFMP: an iterative best basis algorithm for inverse problems in the geosciences*, in Handbook of Geomathematics, W. Freeden, M. Z. Nashed, and T. Sonar, eds., Springer, Berlin, 2015, pp. 2121–2147.

[45] ———, *Geomathematics – Modelling and Solving Mathematical Problems in Geodesy and Geophysics*, Cambridge University Press, Cambridge, 2022.

[46] V. MICHEL AND S. ORZLOWSKI, *On the null space of a class of Fredholm integral equations of the first kind*, J. Inverse Ill-Posed Probl., 24 (2016), pp. 687–710.

[47] ———, *On the convergence theorem for the regularized functional matching pursuit (RFMP) algorithm*, GEM Int. J. Geomath., 8 (2017), pp. 183–190.

[48] V. MICHEL AND N. SCHNEIDER, *A first approach to learning a best basis for gravitational field modelling*, GEM Int. J. Geomath., 11 (2020), Art. 9, 26 pages.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[49] V. MICHEL AND R. TELSCHOW, *A non-linear approximation method on the sphere*, GEM Int. J. Geomath., 5 (2014), pp. 195–224.
[50] ———, *The regularized orthogonal functional matching pursuit for ill-posed inverse problems*, SIAM J. Numer. Anal., 54 (2016), pp. 262–287.
[51] A. F. NIKIFOROV AND V. B. UVAROV, *Special Functions of Mathematical Physics. A Unified Introduction with Applications*, Birkhäuser, Boston, 1988.
[52] Y. C. PATI, R. REZAIIFAR, AND P. S. KRISHNAPRASAD, *Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition*, in Proceedings of 27th Asilomar Conference on Signals, Systems and Computers, vol. 1, Pacific Grove, November 1993, IEEE Computer Society Press, Los Alamitos, pp. 40–44.
[53] R. PLONSEY, *Biomagnetic Phenomena*, McGraw-Hill, New York, 1969.
[54] N. SCHNEIDER, *Learning Dictionaries for Inverse Problems on the Sphere*, PhD. Thesis, University of Siegen, Department of Mathematics, Geomathematics Group, 2020.
[55] N. SCHNEIDER AND V. MICHEL, *A dictionary learning add-on for spherical downward continuation*, J. Geod., 96 (2022), Art. 21, 22 pages.
[56] G. SZEGŐ, *Orthogonal Polynomials*, 4th ed., American Mathematical Society, Providence, 1975.
[57] R. TELSCHOW, *An Orthogonal Matching Pursuit for the Regularization of Spherical Inverse Problems*, PhD. Thesis, University of Siegen, Department of Mathematics, Geomathematics Group, Dr. Hut, Munich, 2014.
[58] P. VINCENT AND Y. BENGIO, *Kernel matching pursuit*, Mach. Learn., 48 (2002), pp. 165–187.
[59] WOLFRAM RESEARCH, INC., *Mathematica*, Version 10.0, 2014.