FULL PAPER

Streaking artifact suppression of quantitative susceptibility mapping reconstructions via L1-norm data fidelity optimization (L1-QSM)

Carlos Milovic[1,2,3,4] [ID] | **Mathias Lambert**[1,2,3] | **Christian Langkammer**[5,6] [ID] |
Kristian Bredies[7,6] [ID] | **Pablo Irarrazaval**[1,2,3,8] | **Cristian Tejos**[1,2,3]

[1]Department of Electrical Engineering, Pontificia Universidad Catolica de Chile, Santiago, Chile

[2]Biomedical Imaging Center, Pontificia Universidad Catolica de Chile, Santiago, Chile

[3]Millennium Nucleus for Cardiovascular Magnetic Resonance, Santiago, Chile

[4]Department of Medical Physics and Biomedical Engineering, University College London, London, UK

[5]Department of Neurology, Medical University of Graz, Graz, Austria

[6]BioTechMed Graz, Graz, Austria

[7]Institute of Mathematics and Scientific Computing, University of Graz, Graz, Austria

[8]Institute for Biological and Medical Engineering, Pontificia Universidad Catolica de Chile, Santiago, Chile

Correspondence
Carlos Milovic, Department of Electrical Engineering, Pontificia Universidad Catolica de Chile, Santiago, Chile.
Email: cmilovic@uc.cl

Funding information
National Agency for Research and Development, ANID, Grant/Award Number: FONDECYT 1191710, Millennium Science Initiative Program – NCN17_12 and PIA-ACT192064; Cancer Research UK, Grant/Award Number: Multidisciplinary Award C53545/A24348; Austrian Science Fund, Grant/Award Number: KLI523 and P30134

Purpose: The presence of dipole-inconsistent data due to substantial noise or artifacts causes streaking artifacts in quantitative susceptibility mapping (QSM) reconstructions. Often used Bayesian approaches rely on regularizers, which in turn yield reduced sharpness. To overcome this problem, we present a novel L1-norm data fidelity approach that is robust with respect to outliers, and therefore prevents streaking artifacts.

Methods: QSM functionals are solved with linear and nonlinear L1-norm data fidelity terms using functional augmentation, and are compared with equivalent L2-norm methods. Algorithms were tested on synthetic data, with phase inconsistencies added to mimic lesions, QSM Challenge 2.0 data, and in vivo brain images with hemorrhages.

Results: The nonlinear L1-norm-based approach achieved the best overall error metric scores and better streaking artifact suppression. Notably, L1-norm methods could reconstruct QSM images without using a brain mask, with similar regularization weights for different data fidelity weighting or masking setups.

Conclusion: The proposed L1-approach provides a robust method to prevent streaking artifacts generated by dipole-inconsistent data, renders brain mask calculation unessential, and opens novel challenging clinical applications such as assessing brain hemorrhages and cortical layers.

KEYWORDS
FANSI, hemorrhage, L1-norm, QSM, total variation

1 | INTRODUCTION

Gradient recalled echo (GRE) sequences encode voxel-wise information about the magnetic fields in the phase of the complex acquisition.[1] The measured field is composed of the contributions of the main field—with its inhomogeneities—plus the magnetic field perturbations introduced by all other objects inside the MR scanner. Such magnetic field perturbations are related to the magnetization of the tissues with different magnetic susceptibilities. For human tissues, the relationship between the magnetic field perturbations and arbitrary susceptibility sources can be modeled as a convolution with the impulse response given by the magnetic dipole kernel.[2,3] Since the appearance of this fast source-to-field method, many researchers have focused their work on solving the inverse problem, ie, finding the susceptibility distribution given the measured magnetic perturbations (field-to-source problem). Direct approaches by solving the system of linear equations are usually infeasible and prone to errors.[4,5] Direct division in the Fourier domain is also discarded due to the ill-posedness of the problem, because the dipole kernel contains a zero-valued double-shaped cone surface (the so-called magic cone) that produces divisions by zero. To avoid this indetermination, truncated kernel strategies were proposed.[6,7] However, these approaches are prone to noise amplification and the appearance of double cone-shaped artifacts (called streaking artifacts). To reduce these effects, regularization strategies were introduced.[8] Most of these strategies formulate the problem as a maximum-likelihood (or data fidelity) term with an additional regularization term that acts as prior knowledge about the solution. The regularizers act as constraints to the solutions, such as imposing continuity or smoothness[9] (classical Tikhonov approaches), piece-wise constant or piece-wise smooth solutions (variational penalties such as total variation [TV][10,11] and total generalized variation[12-16]), sparsity in a specific domain (ie, wavelets[17] or recent low-rank[18] strategies), etc The likelihood of the solution, expressed in the data fidelity term, is typically associated with the L2-norm difference between the acquired phase and the susceptibility distribution convolved with the dipole kernel. From a Bayesian perspective, this is equivalent to assuming that the noise distribution in the phase data is Gaussian. This assumption is only valid for high signal-to-noise ration (SNR) areas, ie, with high magnitude signal.[19] A noise-whitening weight matrix may be used in the data fidelity term to improve the solutions at low-signal regions.[10] Another approach is to change the domain of the data fidelity term and calculate the error in the complex image domain.[20,21] This nonlinear approach is commonly implemented jointly with variational penalties,[20] with proposed fast solvers[22] based on the alternating directions of multipliers method[23-25] (ADMM). These Bayesian approaches improved the robustness against noise and reduced the impact of unwrapping artifacts and the

streaking artifacts. However, strong streaking suppression is usually performed at the expense of reduced sharpness or detail loss.[21,26] In addition, voxels with phase values inconsistent with the "dipole model" will still generate streaking artifacts. Strong inconsistencies may be generated by flow effects, intra-voxel dephasing, and extreme noise, among other sources. Some iterative solutions have been proposed to separate phase contributions that conform to the "dipole model" from other effects and prevent streaking artifact propagation.[27-29] This is also the basis of the dynamic model error reduction[20] (MERIT) algorithm, which updates the weight matrix in the data fidelity term in an adaptive way at each iteration step. All these approaches have in common that they estimate the error using an L2-norm. Due to Parseval's theorem, the least squared error is equivalent to minimizing the energy of the error in the frequency domain. Frequency coefficients of the reconstructions along the magic cone and close to it are usually most affected by noise amplification[30] and by attenuation in over-regularized reconstructions. In the image (susceptibility) domain, this creates a "smearing effect." The phase inconsistencies are propagated following the magic angle and its neighboring voxels to reduce the energy of the errors they create.

In the present work, we present novel data fidelity terms based on estimations of the error with an L1-norm. Instead of minimizing the energy of the errors (least squared error), the L1-norm is equivalent to a least absolute error problem. This promotes a sparser distribution of errors than the L2-norm, preventing energy-spilling. Furthermore, an L1-norm data fidelity term tends to better adapt to non-Gaussian noise distributions, such as impulsive (salt-and-pepper) noise and outliers.[25,31-33] We solve this problem with an ADMM-based approach and compare the performance of these new L1-norm methods against their equivalent L2-norm versions in the phase (linear problem) and complex image domains (nonlinear problem).

2 | METHODS

The phase of GRE acquisitions Φ is proportional to the magnetic field perturbations of the main field[1-3]:

$$\phi = \gamma B_0 TE \cdot F^H \left(\frac{1}{3} - \frac{k_z^2}{k^2} \right) F \chi = F^H D F \chi, \qquad (1)$$

where F is the Fourier Transform and F^H its inverse, B_0 is the main field strength, γ is the gyromagnetic ratio, TE is the echo time of the acquisition, k_z the frequency domain index in the z-direction and $k^2 = k_x^2 + k_y^2 + k_z^2$. D becomes the dipole kernel in the frequency domain (voxel-wise multiplication), and χ is the susceptibility of the tissues (or a convolution in the spatial

domain between χ and the dipole unit). We assume that Φ is the local phase, after background field removal, and removal of the offset through phase evolution estimation via multi-echo fitting. Flow, chemical shift, and fat-water-associated effects are ignored in this model.

Most QSM algorithms reformulate the problem as an optimization problem with two or more additive terms[26]:

$$\text{argmin}_\chi \left\| W(F^H DF \chi - \phi) \right\|_2^2 + \alpha \cdot \Omega(\chi), \qquad (2)$$

where the data fidelity term is a measure of the error between the phase calculated from the estimated χ and the acquired Φ data, with W a voxel-wise weight (defined by a region-of-interest (ROI) binary mask, or a weight proportional to the signal-to-noise-ratio (SNR) or signal magnitude). $\Omega(\chi)$ is an additional regularizer. The two terms are balanced by a Lagrangian multiplier (or regularization weight), $\alpha > 0$. Typical regularizers include the classical Tikhonov term,[8] TV[9,34], and other variational penalties.[10,14] Optimization of Equation (2) using TV as regularizer is hereafter named the "linear L2-norm" method.

From a Bayesian perspective, the data-fidelity term used in Equation (2) models Gaussian noise in the acquired phase data. Unfortunately, the noise distribution is Gaussian in the acquired complex signal, but not in the signal phase.[19] This leads to reconstruction errors in low SNR regions and streaking artifacts. Instead of modeling the non-Gaussianity of the noise distribution in the signal phase, a regularized solution of a nonlinear forward model was proposed.[20] This functional projects the phase data onto the complex signal domain (with W typically the signal magnitude):

$$\text{argmin}_\chi \left\| W \left(e^{iF^H DF \chi} - e^{i\phi} \right) \right\|_2^2 + \alpha \cdot \Omega(\chi). \qquad (3)$$

Optimization of Equation (3) using TV as regularizer is hereafter named the "nonlinear L2-norm" method. This functional is also robust against some unwrapping errors. Background field residuals may also propagate as errors into the reconstruction, usually in the form of low-frequency features. These errors may be reduced by incorporating an additional phase component to the data fidelity term and forcing this component to be harmonic by a second regularizer (weak-harmonicity constraint).[35] Other errors in the phase, such as coil combination errors, intravoxel and flow-related dephasing, echo combination errors, could still be a source of strong streaking artifacts. The effects of some of those errors may be reduced using L1-norm data fidelity terms. Although from a Bayesian point of view an L1-norm data fidelity term models a Laplacian noise, this is more robust than non-Gaussian noise distributions (for example, Poisson noise) and outliers.[25,31-33]

3 | L1-NORM OPTIMIZATION

We present L1-type QSM solvers for both the linear and non-linear problems.

3.1 | Linear L1-norm solver

The linear L1-norm optimization problem is defined as:

$$\text{argmin}_\chi \left\| W(F^H DF \chi - \phi) \right\|_1 + \alpha \cdot TV(\chi), \qquad (4)$$

where the anisotropic $TV(\chi) = \nabla \chi_1$ regularizer is used for simplicity, although the isotropic alternative and extension to Total Generalized Variation are feasible.[16,22] We solve this optimization problem using the ADMM.[23-25,36] In the ADMM framework, additional auxiliary (or splitting) variables are introduced. These variables substitute the original variable to be optimized, or functions of it, and equality is imposed by new constraints reintroduced as penalties and Lagrange multipliers. The augmented functional is minimized for each variable (primal and auxiliary) while fixing the others. These subproblems may be easier to solve, often with closed-form solutions. To complete one iteration step, all subproblems are solved, and all Lagrange multipliers are updated. Iterations are stopped when the original variables reach convergence, indicated by an appropriate criterion. In our Equation (4), an auxiliary variable $z = F^H DF \chi - \phi$ is needed to decouple the data fidelity term from the regularization term. Similarly, the splitting variable $z_{TV} = \nabla \chi$ is used to deal with the L1-norm term in the TV regularizer. The augmented problem becomes:

$$\text{argmin}_{\chi,z} \|Wz\|_1 + \frac{\mu}{2} \left\| F^H DF \chi - \phi - z + s \right\|_2^2 \\ + \alpha \cdot \|z_{TV}\|_1 + \frac{\mu_{TV}}{2} \left\| \nabla_\chi - z_{TV} + s_{TV} \right\|_2^2. \qquad (5)$$

In this framework, μ, $\mu_{TV} > 0$ are scalar Lagrangian weights, with s and s_{TV} Lagrangian multipliers of the same size as z and z_{TV}. The χ and z_{TV} subproblems are solved as previously reported[16,22] with closed-form solutions, described in the Supporting Information, which is available online.

The z subproblem is solved using a proximal operation:

$$z = \max \left(\left| F^H DF \chi - \phi + s \right| - \frac{W}{\mu}, 0 \right) \cdot \text{sign} \left(F^H DF \chi - \phi + s \right). \qquad (6)$$

For each iteration, the solutions for the χ, z_{TV} and z sub-problems are calculated sequentially. At the end of each (n) iteration, the Lagrange multipliers s_{TV} and s are updated (see the Supporting Information for more details on these updates). The update of s is given by:

$$s \leftarrow F^H D F \chi - \phi - z + s. \qquad (7)$$

3.2 | Nonlinear L1-norm solver

The analogous nonlinear[20-22] L1-norm optimization problem is:

$$\operatorname{argmin}_\chi \left\| W \left(e^{iF^H DF\chi} - e^{i\phi} \right) \right\|_1 + \alpha \cdot TV(\chi). \qquad (8)$$

To minimize this functional, two auxiliary variables are introduced: $z_1 = F^H DF\chi$ and $z_2 = e^{iz_1} - e^{i\phi}$, in addition to z_{TV}. The augmented problem becomes:

$$\operatorname{argmin}_{\chi,z_1,z_2} \frac{1}{2} \|Wz_2\|_1 + \frac{\mu_1}{2} \left\| F^H DF\chi - z_1 + s_1 \right\|_2^2 + \frac{\mu_2}{2} \left\|$$
$$e^{iz_1} - e^{i\phi} - z_2 + s_2 \right\|_2^2 + \alpha \cdot \|z_{TV}\|_1 + \frac{\mu_{TV}}{2} \left\| \nabla_\chi - z_{TV} + s_{TV} \right\|_2^2. \qquad (9)$$

e χ and z_{TV} subproblems are thus solved in the same augmentation way as in Equation (5). The z_2 subproblem is again solved with a proximal operation:

$$z_2 = \max \left(\left| e^{iz_1} - e^{i\phi} + s_2 \right| - \frac{W}{\mu_2}, 0 \right) \cdot \operatorname{sign} \left(e^{iz_1} - e^{i\phi} + s_2 \right). \qquad (10)$$

The z_1 subproblem is solved as in the Fast Nonlinear Susceptibility Inversion (FANSI)[22] algorithm, with a Newton-Raphson iterative approach:

$$z_{1,k+1} = z_{1,k} + \frac{\mu_2 \sin\left(z_{1,k} - angle\left(e^{i\phi} + z_2 - s_2\right) - i \cdot log\left(\left|e^{i\phi} + z_2 - s_2\right|\right)\right) + \mu_1 z_{1,k} - \mu_1\left(F^H DF\chi + s_1\right)}{\left(\mu_2 \cos\left(z_{1,k} - angle\left(e^{i\phi} + z_2 - s_2\right) - i \cdot log\left(\left|e^{i\phi} + z_2 - s_2\right|\right)\right) + \mu_1\right)}. \qquad (11)$$

This inner iterative solver stops when a normalized update of 10^{-6} (or lower) is achieved or at a maximum of 10 iterations, whichever is satisfied first, and is initialized with:

$$z_{1,0} = F^H DF\chi + s_1. \qquad (12)$$

Finally, the updates for s_1 and s_2 are given by:

$$s_1 \leftarrow F^H DF\chi - z_1 + s_1, \qquad (13)$$

$$s_2 \leftarrow e^{iz_1 - e^{i\phi}} - z_2 + s_2. \qquad (14)$$

In all of our experiments we used $\mu = \mu_1 = \mu_2 = 1.0$. Since the z and z_2 updates for the solution of Equation (5) and Equation (9) require a proximal operation involving W, we

analyzed how a positive scalar factor, λ, affected the results. We tested the following weighting methods: (a) no Mask: $W = \lambda$, (b) ROI mask: $W = \lambda \cdot mask$ and (c) magnitude weight: $W = \frac{\lambda \cdot mask \cdot Magn}{max(Magn)}$. Here, $mask$ is a binary volume defined by a ROI that selects only the brain. We used $\mu_{TV} = 100\alpha$ as suggested by previous reports.[22,35] The ADMM outer iterations are set to stop after a fixed number of iterations (default = 300) or after a given update percentage, $100 \cdot \frac{\chi_k - \chi_{k-1}}{\chi_{k-1}}$, is reached (default = 0.1%). The condition that is satisfied first stops the solver.

The proposed numerical solvers were implemented in MATLAB 2019a (The Mathworks Inc, USA) using an Intel i7 9750H processor (@4.5GHz/32GB RAM) and an Intel i7-2600 (@3.40GHz/32GB RAM). These new algorithms were compared against linear and nonlinear solvers with L2-norm data fidelity terms, as included in the FANSI[22] Toolbox (also similar[22] to other state-of-the-art methods[10,20]).

4 | EXPERIMENTAL DESIGN

4.1 | COSMOS-brain simulation

We used the Calculation Of Susceptibility through Multiple Orientation Sampling (COSMOS)[37] reconstruction included in the 2016 QSM Reconstruction Challenge (RC1)[38] dataset (available at http://qsm.neuroimaging.at) as susceptibility ground truth. We forward-simulated the phase and added complex Gaussian noise (using the provided magnitude data), with peak SNR = 100. This was used as input for the QSM reconstructions. We optimized the free reconstruction parameters (λ and α) for all three data-fidelity weighting methods (described in the previous section) by minimization of the normalized root mean square of error (RMSE, given in percentages, normalized by the L2-norm of the ground truth) and maximization of the Structural Similarity Index Metric[39-41] (SSIM). To optimize λ and α, we set the stopping rule to the default values. For each optimal set of parameters found by this procedure, we measured the convergence rate (update percentage) and evolution of the error metrics up to 2500 iterations.

The λ and α parameters and weighting method performing best for each algorithm were chosen to reconstruct additional simulations, where the phase data were modified in the following manner:

4.1.1 | Phase jump test

Five phase single-voxel jumps were introduced to generate dipole inconsistencies. Their values were chosen as -27π, -13.5π, 6.75π, 13.5π, and 27π. The purpose was to test the algorithm's response to localized large phase inconsistencies. These phase jumps were placed in the central XY plane at randomly selected locations inside the brain.

4.1.2 | Extrapolation test

To test the algorithm's response to extended phase inconsistencies, two spheres of radius six voxels representing diamagnetic (-0.3 ppm) and paramagnetic (0.15 ppm) concentrations were added to the COSMOS ground truth. The spheres were randomly located along the z central axis within the brain, assuming strong dephasing inside them (zero value in the magnitude image). For simplicity, no extralesional signal dephasing is considered in this simulation. We compared reconstructions with and without masking the spheres (updated ROI mask, with $mask_{spheres} = 0$ inside the spheres, or the same $mask$ as in (a), respectively).

4.1.3 | Susceptibility jump test

Five single-voxel susceptibility jumps (-0.5, -0.25, 0.1, 0.25, and 0.5 ppm) were introduced to the COSMOS ground truth to test the algorithm's response to high contrast susceptibility sources. They were placed in the same XY plane as in (a), with transposed X and Y coordinates.

4.2 | SIM2SNR1 data

The 2019 QSM Reconstruction Challenge (RC2 - Seoul, Korea) provided two simulated multi-echo GRE acquisitions (Sim1 and Sim2), with known ground-truth susceptibility maps.[42-44] Sim2 has greater contrast between white and gray matter than Sim1, and it also contained a simulated calcification. Since the susceptibility simulation was performed at a higher resolution than the simulated phases, strong dephasing effects are present at the calcification and vessels. We used the supplied field map for the lowest SNR (SNR1) to make the evaluation more challenging. Optimal parameter settings were obtained to minimize global RMSE. We calculated all the error metrics used for the challenge competition[44]: RMSE, RMSE detrended (dRMSE), RMSE in specific ROIs: Blood (dRMSE_Blood), Tissues (dRMSE_Tissue), and Deep Gray Matter (dRMSE_DGM), Calcification Streaking (CalcStreak), and deviation from calcification moment (DFCM), plus additional global metrics such as SSIM,

high frequency error norm[38] (HFEN), correlation coefficient (CC), mutual information (MI) and the normalized mean absolute error (MAE, given in percentages, normalized by the L1-norm of the ground truth). To optimize λ and α we used the default stopping rule as described for the COSMOS-brain example. These optimal parameters were subsequently used to analyze the evolution of the metric scores up to 2500 iterations (at a 10-iteration interval). Visualizing the results and calculating the error metrics were done using the source code supplied by the QSM Challenge committee in the Report paper.[44]

4.3 | In vivo data

We performed an in vivo acquisition on a Siemens 3T scanner (Magnetom Trio Tim; Siemens Healthcare, Erlangen, Germany) with a 12-channel phased-array head coil. GRE sequence with six echoes. A patient showing extensive hemorrhage in the brain was scanned with the following sequence parameters: $TE_1 = 4.92$ ms, $\Delta TE = 4.92$ ms, TR = 35 ms, flip angle = 15°, $232 \times 288 \times 64$ matrix with $0.8 \times 0.8 \times 2$ mm^3 voxel size, and $T_{acq} = 4{:}51$ min. Preprocessing included nonlinear multi-echo combination[20] and background field removal performed by the projection onto dipole fields method (PDF).[45] Background field residuals were removed using the harmonic phase estimation obtained with the weak-harmonic QSM method (WH-TV)[35] and generated a new corrected local field map. MRI data were acquired with the approval of the local institutional review board.

In addition, we used the provided single orientation 3T (Tim Trio, Siemens) in vivo acquisition from the RC1[38] dataset to test our algorithms. TE/TR = 25/35 ms, and $T_{acq} = 92$s (15-fold acceleration) $160 \times 160 \times 160$ matrix with 1.0 mm^3 isotropic voxels. Background field removal was performed by the Laplacian Boundary Value method[46] followed by 4th order polynomial fit.[38]

Optimal reconstruction parameters were found by an initial estimation using the L-curve analysis,[11] followed by visual fine-tuning to produce equivalent results for all algorithms.

5 | RESULTS

5.1 | COSMOS-brain simulation

Parameter optimization was performed for all three data fidelity weighting methods using RMSE (Figure 1) and SSIM. Due to instabilities in the Newton-Raphson inner solver, the nonlinear L2-norm algorithm was unable to achieve results without a brain mask ("no mask"). Optimal SSIM reconstructions were achieved with similar parameters in all cases, except for the linear L2-norm method, which provided
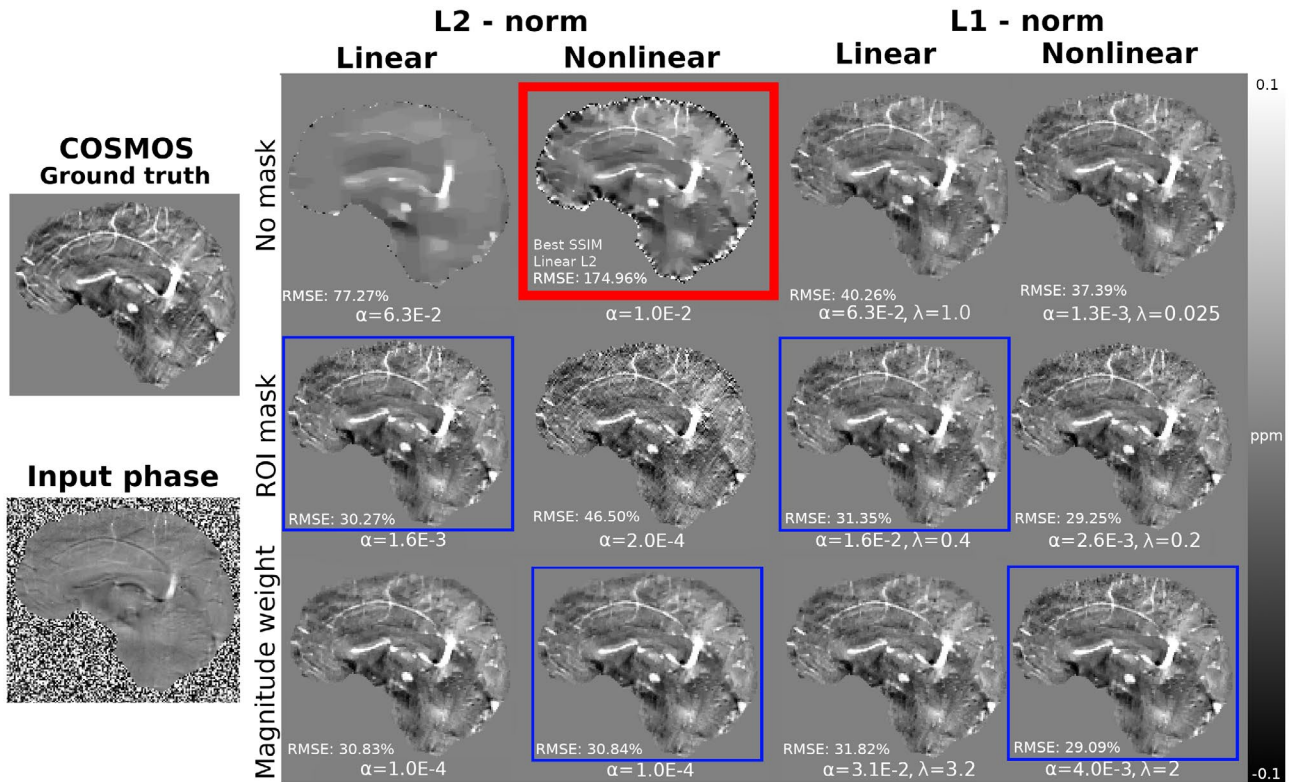
FIGURE 1 Reconstructions of forward-simulated noise-corrupted (peak SNR = 100) COSMOS data. Each column presents the results for a different data fidelity term. Rows show the results of using different data fidelity weighting strategies (no mask, ROI mask, and magnitude weighting). RMSE scores and optimal weighting parameters (α and λ) are also presented, calculated inside the ROI. Blue squares mark the results with the lowest RMSE for each algorithm. Please note that the nonlinear L2-norm algorithm diverged when no mask was used in the data fidelity term, and the best SSIM solution using the linear L2-norm algorithm is presented in its place (highlighted with a red box). All reconstructions are shown masked for displaying purposes (external noisy voxels set to zero). The same input phase was used in all cases (left). For all L2-norm methods we used $\lambda = 1$

optimal parameters considerably smaller than those obtained optimizing RMSE. In this case, the SSIM-optimized reconstruction (shown with a red square in Figure 1) yielded more anatomical details than the RMSE-optimized reconstruction, although with large errors in the boundary (absolute errors shown in Supporting Information Figure S1). Despite the chosen error metric, the linear L2-norm method returned over-regularized results. L1-norm results for the "no mask" method contained more anatomical details than L2-norm algorithms, lacking noticeable errors at the boundary. Using either an "ROI mask" or "Magnitude weight" improved the reconstruction quality for both L1-norm and L2-norm methods (Supporting Information Table S1). Visually, the largest structural differences between "ROI mask" and "Magnitude weight" were found in cortical areas, where "Magnitude weight" results tended to over-smooth solutions and loose definition of the veins and other small details (see yellow arrows in Supporting Information Figure S1). In this experiment, the best results were achieved with "ROI mask" for the linear methods and "Magnitude weight" for the nonlinear ones. The nonlinear L1-norm method achieved overall the lowest RMSE and largest SSIM scores. MAE scores were

highly correlated with RMSE. The best results for each method are also shown in other orientations in Supporting Information Figure S2. L2-norm methods achieved optimal results at $\lambda = 1.0$, but no relevant differences were obtained when changing λ values around 1.0. L1-norm methods showed a larger sensitivity to λ. Large λ values may also cause the nonlinear L1-norm method to diverge.

All algorithms have a quasi-monotonic behavior in RMSE and SSIM, converging after approximately 100 iterations (Figure 2). A slight degradation is produced for large numbers of iterations (Supporting Information Table S2). In most cases, a 0.1% update correlated well with optimal error metrics. The nonlinear L2-norm method required more iterations to achieve optimal results, closer to a 0.01% update. The nonlinear L1-norm method seemed to be unstable in terms of the update rate for many iterations, although the error metrics showed little variation. Both the linear L2-norm and the nonlinear L1-norm methods showed faster improvement in error metrics for early iteration stages. Visually, most of the differences between the results at a 0.1% update and the final 2500 iterations were restricted to the vessels, as shown in Supporting Information Figure S3. This also revealed

substantially larger errors for the nonlinear L2-norm method in large paramagnetic areas. Errors in the vessels were still present for all methods at 2500 iterations, again more noticeable for the nonlinear L2-norm method. In terms of executing times, the linear L1-norm method presented a similar per-iteration speed as its L2-norm counterpart, whereas the nonlinear L1-norm was up to 45% slower than the linear L1-norm method.

When single-voxel phase inconsistencies were introduced, the linear L2-norm solution presented strong streaking artifacts (Figure 3). The nonlinear L2-norm method mitigated this effect due to its robustness against 2π jumps. However, artifacts are still noticeable in the neighborhood of the phase jumps. Both L1-norm methods rejected the inconsistencies resulting in virtually streaking artifact-free reconstructions.

Figure 4 shows the results for the extrapolation test using two spheres. When inconsistencies are included in the ROI (*Unmasked* results, using the original *mask* instead of *mask-spheres*), both L2-norm methods propagated the errors into streaking artifacts. Both L1-norm methods, on the contrary,
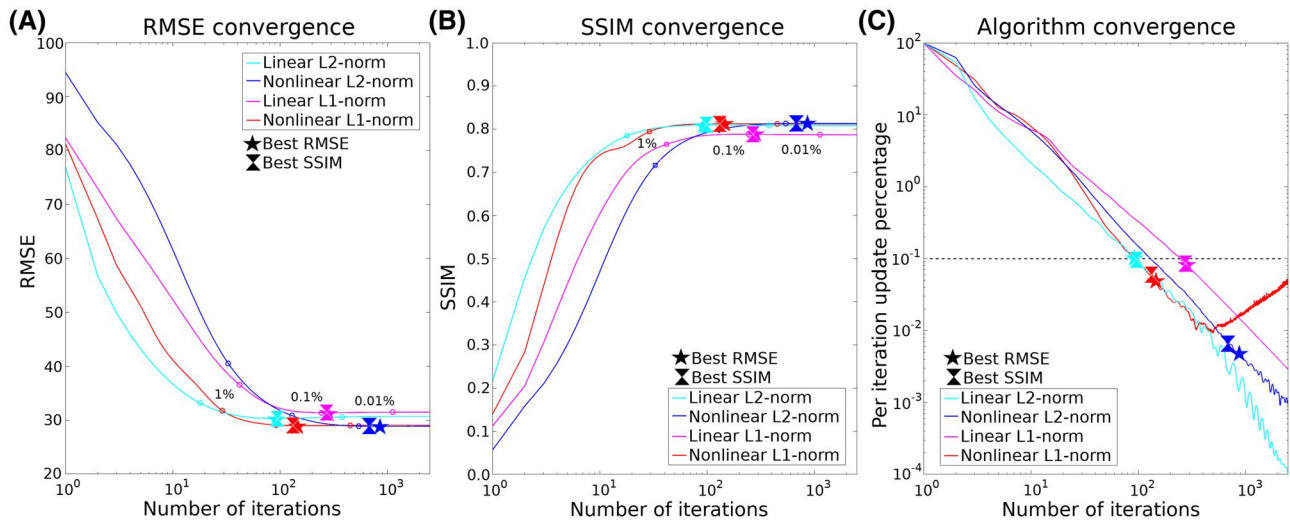

FIGURE 2 Evolution of the RMSE (A), SSIM (B), and algorithm convergence (C) (measured by the normalized change in the susceptibility map between each iteration) up to 2500 iterations, in the COSMOS-based forward simulation experiment
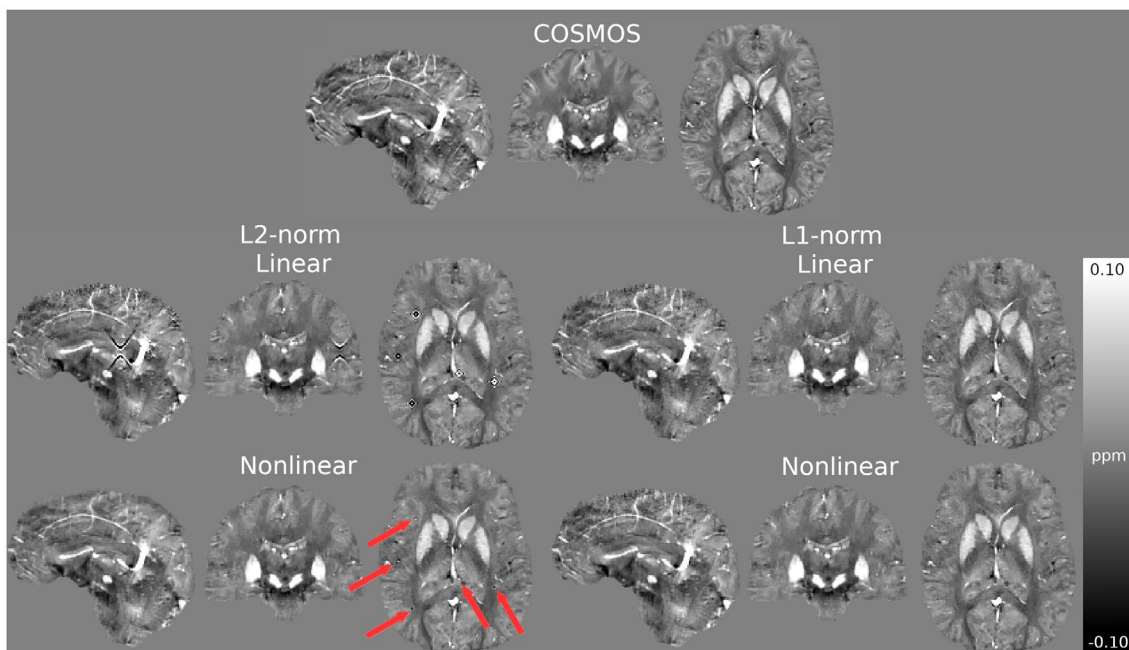

FIGURE 3 Reconstructions of COSMOS-based forward simulations degraded with 5 single voxel large phase inconsistencies. Same weighting parameters used as in the ground-state (not degraded) simulation. Red arrows indicate the location of the single-voxel phase jumps. Note the presence of streaking artifacts in coronal and sagittal slices, visible as ringing in axial slices in L2-norm methods
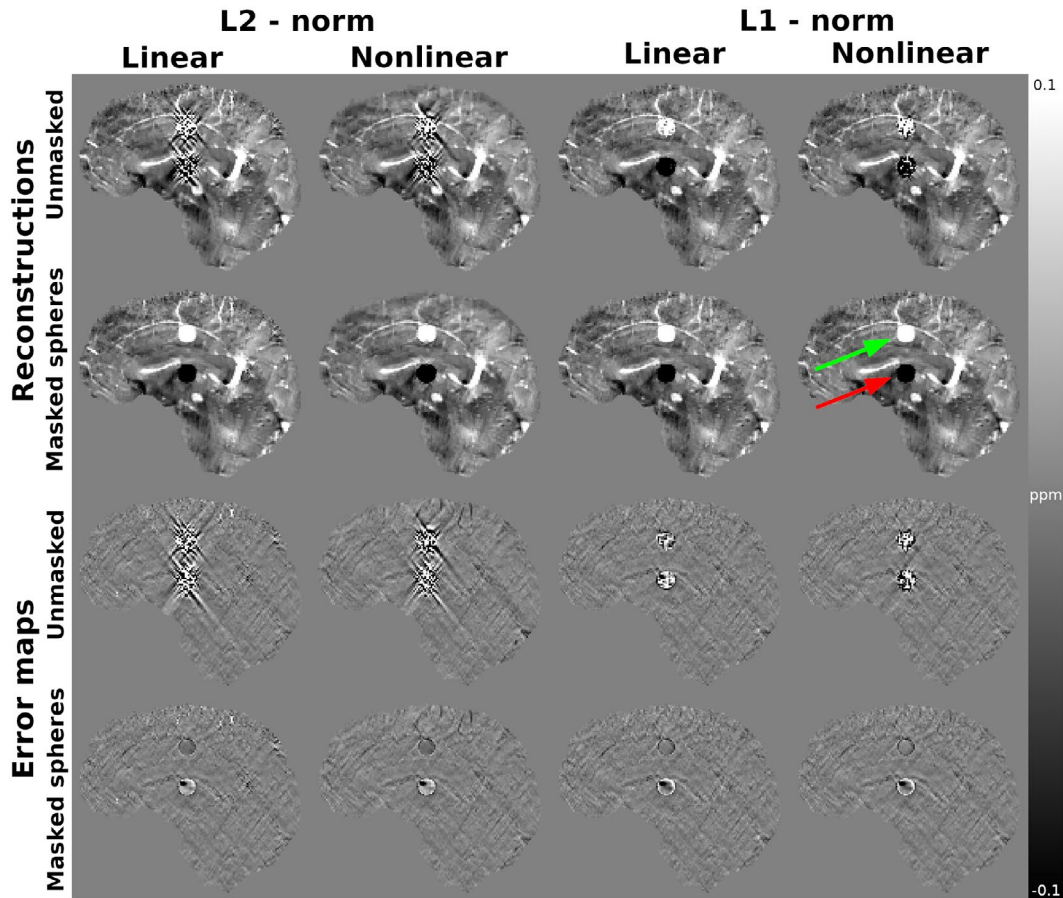
FIGURE 4 Reconstructions of COSMOS-based forward simulations which included two spheres simulating diamagnetic (red arrow, −0.30 ppm) and paramagnetic (green arrow, +0.15 ppm) tissues, and zero effective signal magnitude. Results are provided with and without masking of the spheres (data fidelity weighting). The latter provides an extrapolation test, while the first tests the robustness against signal inconsistencies. Same weighting parameters used as in the ground-state (not degraded) simulation. Error maps show the difference between the reconstruction and the ground truth

limited the extent of the errors within the discrepancy region. Quantitatively, mean values inside the spheres were also reconstructed more accurately by the L1-norm methods than by the L2-norm methods (from 10% to 40% local RMSE for a 0.1% update as stopping criteria). As shown in Supporting Information Figure S4, all methods diverged after a few iterations in global RMSE, whereas SSIM showed a similar behavior as for the ground-state case. Global RMSE showed at least 15% improvements for both L1-norm methods over the L2-norm methods. Results for *Masked spheres* (using *mask-spheres*) are in practice inpainting the contents of the spheres (ie, inside the spheres, all the information is replaced or filled following the external phase data and is strongly regularized). In this case, L1-norm methods were also more accurate.

Finally, experiments with local strong susceptibility sources are shown in Supporting Information Figure S5. All methods obtained similar results, both visually and in terms of relative errors. The susceptibility source of 0.05 ppm, located contiguous to a vessel, presented the largest relative errors, with the linear L2-norm method being considerably worse (65%) than the other methods (between 22% and 36%).

Global error metrics for all these experiments are presented in Table 1. Please note that normalization of the metrics was performed using the respective ground-truth for each experiment. This included the embedded spheres or strong susceptibility sources, yielding lower relative errors than the discrepancy-free experiment.

5.2 | SIM2SNR1 data

Optimal reconstructions of the RC2 data are presented in Figure 5. Detailed metric scores and reconstruction parameters are presented in Table 2 for different numbers of iterations. Overall, the nonlinear L1-norm method presented the lowest RMSE, HFEN, and MAE errors, the highest scores for the SSIM and MI metrics, reduced streaking propagation (CalcStreak), and the most accurate representation of the calcification (DFCM).

In contrast to the previous COSMOS-simulated experiment, the evolution of the SSIM metric showed a large degradation with a large number of iterations (Figure 6), whereas

RMSE-based and calcification-dependent scores benefited from more iterations. The linear L1-norm method needed more than 2500 iterations to reach proper convergence in the RMSE-based and Calcification-based metrics (CalcStreak and DFCM). As for the previous experiment, errors at the vessels were larger than for other structures, as reflected in dRMSE_Blood.

Results with a 0.1% update stopping criterion are presented in Supporting Information Figure S6. Some minor changes are visible at the calcification and streaking propagation,

TABLE 1 SSIM and RMSE metric scores of reconstructions for all the COSMOS-based forward simulated experiments

| | | 0.1% update | | | | Best reconstruction | | | |
| | | L2-norm | | L1-norm | | L2-norm | | L1-norm | |
Stopping rule:		Linear	Nonlinear	Linear	Nonlinear	Linear	Nonlinear	Linear	Nonlinear
Ground-state	SSIM	0.810	0.799	0.788	0.812	0.810	0.813	0.788	0.813
	RMSE	30.3	30.8	31.4	29.1	30.3	28.8	31.3	29.0
Phase errors (single-voxel)	SSIM	0.759	0.798	0.788	0.812	0.771	0.798	0.788	0.812
	RMSE	143.8	31.1	31.4	29.1	52.3	31.1	31.4	29.1
Spheres (unmasked)	SSIM	0.779	0.783	0.785	0.804	0.783	0.785	0.785	0.806
	RMSE	81.5	75.8	33.7	46.6	46.3	50.0	33.0	33.4
Masked spheres	SSIM	0.808	0.795	0.787	0.811	0.809	0.812	0.788	0.812
	RMSE	30.2	31.8	31.1	29.1	30.2	28.8	31.0	28.8
Susc. Sources (single-voxel)	SSIM	0.810	0.799	0.788	0.812	0.810	0.813	0.788	0.813
	RMSE	30.3	30.8	31.4	29.1	30.3	28.8	31.3	29.0

Notes: All experiments shared the same reconstruction parameters for each algorithm. Metric scores are shown for a stopping criterion of a 0.1% update, and the best scores up to 2500 iterations. Note that the ground truth was modified for the experiments containing two spheres. Due to the large absolute susceptibilities of theses spheres, RMSE scores are normalized by a larger factor than for the other experiments. This resulted in lower RMSE scores for the *Masked spheres* case than for the Ground-state experiment (without spheres).
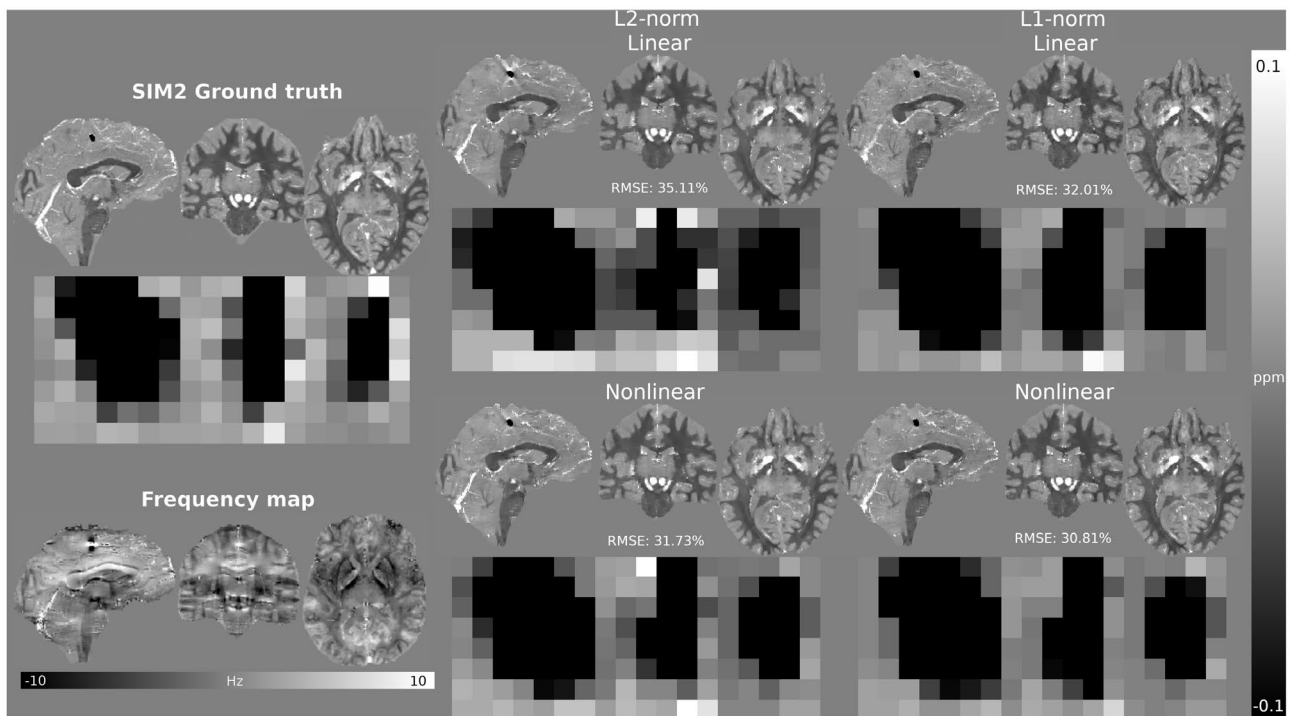

FIGURE 5 Optimal (RMSE-based) reconstructions of the RC2 SIM2SNR1 dataset. The provided frequency map was used as input (left), scaled to radians with a simulated TE of 10 ms to minimize streaking in the L2-norm methods. A magnified view around the calcification is provided for each algorithm. Reconstructions were masked in accordance to the evaluation ROI (the ground truth contains two additional voxels at the boundaries in all directions)

TABLE 2 Metric scores for the RC2 SIM2SNR1 reconstructions are presented, for different number of iterations

Stopping rule:	0.1% update				Best RMSE				2500 iterations			
	L2-norm		L1-norm		L2-norm		L1-norm		L2-norm		L1-norm	
	Linear	Nonlinear	Linear	Nonlinear	Linear	Nonlinear	Linear	Nonlinear	Linear	Nonlinear	Linear	Nonlinear
RMSE	35.4	31.8	34.2	31.8	35.1	31.7	32.0	30.8	35.2	32.3	32.0	30.8
dRMSE	35.4	31.7	34.2	31.8	35.1	31.7	32.0	30.8	35.2	32.3	32.0	30.8
dRMSE_Tissue	37.4	32.9	35.0	32.4	37.3	32.8	33.4	32.0	37.5	33.8	33.4	32.0
dRMSE_Blood	66.3	59.3	63.0	60.8	65.7	59.4	56.9	56.8	65.7	58.7	56.9	56.8
dRMSE_DGM	24.0	24.3	24.0	24.6	24.0	24.3	24.2	24.4	23.8	24.0	24.2	24.4
DevLinearSlope	1.0E-02	6.6E-03	-3.1E-03	4.5E-03	6.8E-03	6.9E-03	-3.4E-03	-1.9E-03	3.4E-03	-5.1E-04	-3.4E-03	-2.4E-03
offset	-1.1E-10	5.3E-11	1.5E-10	-1.9E-09	4.2E-10	-1.5E-09	1.0E-10	8.8E-10	-9.8E-10	1.1E-09	-1.2E-10	-5.2E-10
CalcStreak	0.061	0.024	0.035	0.027	0.048	0.025	0.021	0.017	0.040	0.017	0.021	0.016
DFCM	25.0	12.8	13.5	11.4	22.4	13.0	10.8	8.2	20.3	10.4	10.8	8.0
SSIM	0.782	0.801	0.762	0.808	0.764	0.802	0.765	0.776	0.746	0.761	0.765	0.775
HFEN	34.2	28.0	30.3	27.8	33.0	28.0	27.7	26.0	32.2	27.6	27.7	26.0
CC	0.931	0.945	0.930	0.944	0.929	0.945	0.938	0.941	0.925	0.936	0.938	0.941
MI	0.493	0.508	0.516	0.527	0.435	0.510	0.506	0.456	0.398	0.464	0.506	0.444
MAE	29.9	28.4	32.3	28.1	31.5	28.3	32.0	31.3	33.2	32.4	32.0	31.4
Iterations	254	253	361	266	520	240	2500	1520	2500	2500	2500	2500
α (Re. weight)	2.0E-03	1.6E-03	6.8E-02	4.8E-03	2.0E-03	1.6E-03	6.8E-02	4.8E-03	2.0E-03	1.6E-03	6.8E-02	4.8E-03
λ (W = λMagn)	1	1	1.91	0.32	1	1	1.91	0.32	1	1	1.91	0.32

Notes: This considers reconstructions stopped when a 0.1% update was achieved, and at 2500 iterations. Scores for the reconstructions stopped when the best RMSE was achieved are also included (other metric scores were not considered). Overall best scores for each metric are presented highlighted in green. Demeaning and detrending scores reveal no relevant attenuation or offsets for any method compared. Reconstruction parameters and effective number of iterations are also reported.
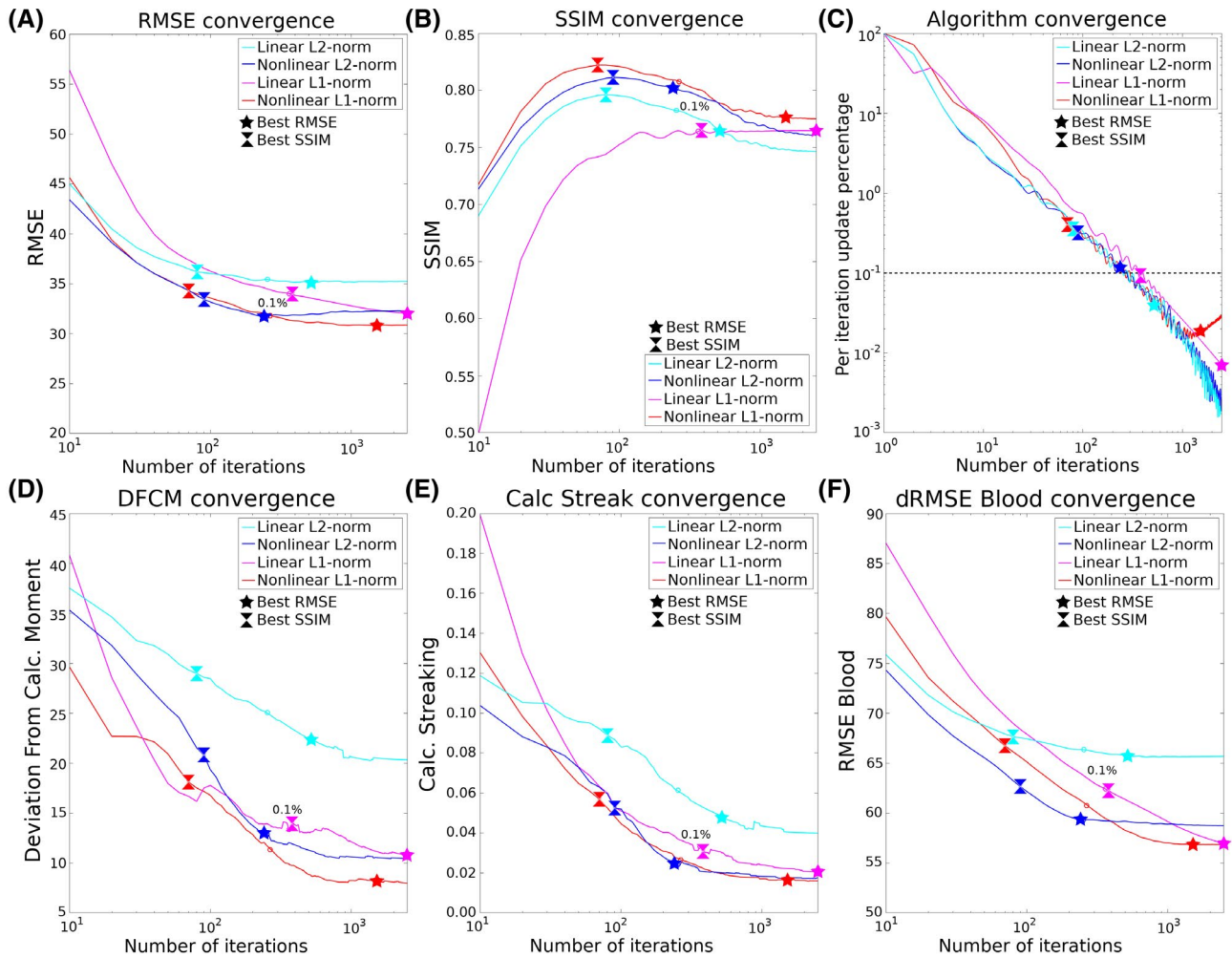
FIGURE 6 Evolution of the RMSE, SSIM, algorithm convergence, deviation from calcification moment (DFCM), calcification streaking (CalcStreak), and RMSE of the blood-related regions (dRMSE_Blood) of reconstructions using the RC2 SIM2SNR1 dataset. Metrics were calculated every 10 iterations, up to 2500 iterations

compared to the optimal results shown in Figure 5. The results with the final 2500 iterations are not shown, as they are visually indistinguishable from those presented in Figure 5.

5.3 | In vivo data

In Figure 7 are presented the preprocessing steps followed to generate the corrected local field map used as input for the in vivo QSM recontruction of a patient with a brain hemorrhage. Due to background field residuals in the local field calculated using PDF, we used the WH-TV method to estimate these residuals.[35] The estimated residual field was subtracted from the PDF result and masked. Figure 8 shows optimal reconstructions using the corrected local field map according to the L-curve analysis[30] with further visual fine-tuning to achieve similar reconstructions for all methods. Both L2-norm methods showed strong streaking artifacts around the lesions (see frontal and posterior lesions)

and veins. In contrast, both L1-norm methods efficiently suppressed these artifacts (although they do not completely remove extralesional contributions). As for the previous experiments, the linear L1-norm produced a noisier image compared to all other methods, while the nonlinear L1-norm method produced smoother images, similar to those obtained with the L2-norm methods.

Results for the RC1 in vivo dataset are presented in Supporting Information Figure S7. This experiment also revealed similar results for both L2-norm methods. L1-norm results differed notably for the largest vessel (great vein of Galen), where the nonlinear method seemed to underestimate its susceptibility values. Still, it prevented the streaking propagation originated at this vessel. Extravascular contributions were reduced for L1-norm methods. Both L1-norm approaches also showed better gray/white contrast in cortical areas compared to L2-norm methods. For comparison, the susceptibility ground truth[47] built from a projection of the tensorial susceptibility components[48]
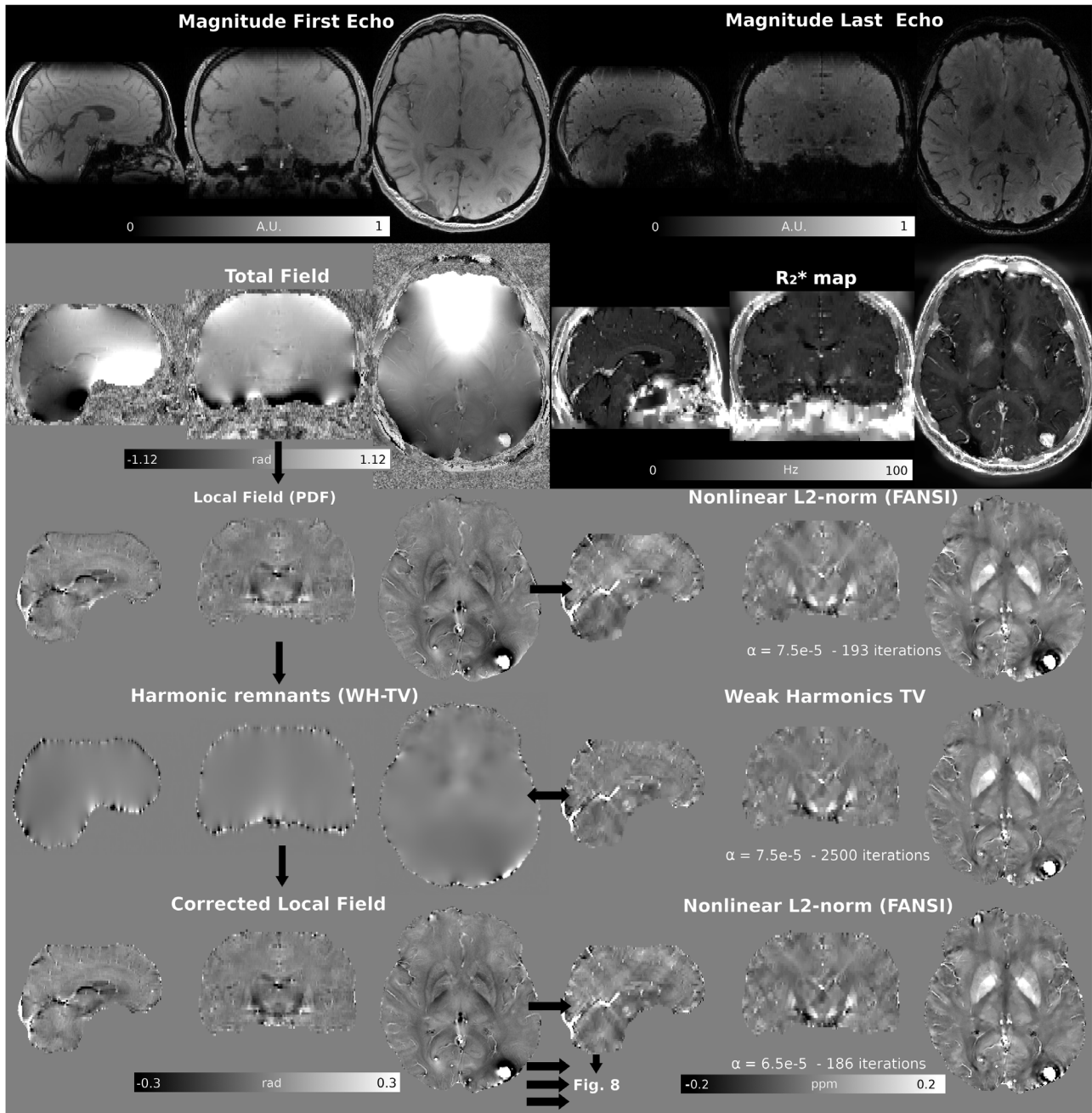
FIGURE 7 Pre-processing pipeline for the in vivo dataset of a patient with a brain hemorrhage. The total field map was estimated using nonlinear multi-echo fitting and scaled back to radians with the echo spacing time. A brain mask was generated using the R2* and the magnitude of the first echo (top). The local field generated by the PDF method still contained background field residuals, which were estimated using the WH-TV algorithm. Harmonic residuals were removed from the local field, and the resulting map was used as input for the final reconstructions shown in Figure 8

is presented. Compared to L2-norm methods, L1-norm methods presented a lower discrepancy to this ground truth in cortical areas. This susceptibility tensor-derived ground truth was calculated using the complete RC1 multi-orientation dataset. The resulting SNR and contrast are higher than for reconstructions performed with the provided highly accelerated and noisy single-orientation acquisition.

6 | DISCUSSION

QSM is an ill-posed inverse problem where the dipole kernel is a nonlocal operator with a zero-valued surface (the magic cone) that spans both high- and low-frequency components. Traditional methods to solve this problem showed that small-sized but strong phase inconsistencies propagate through the magic cone, degrading susceptibility maps with streaking
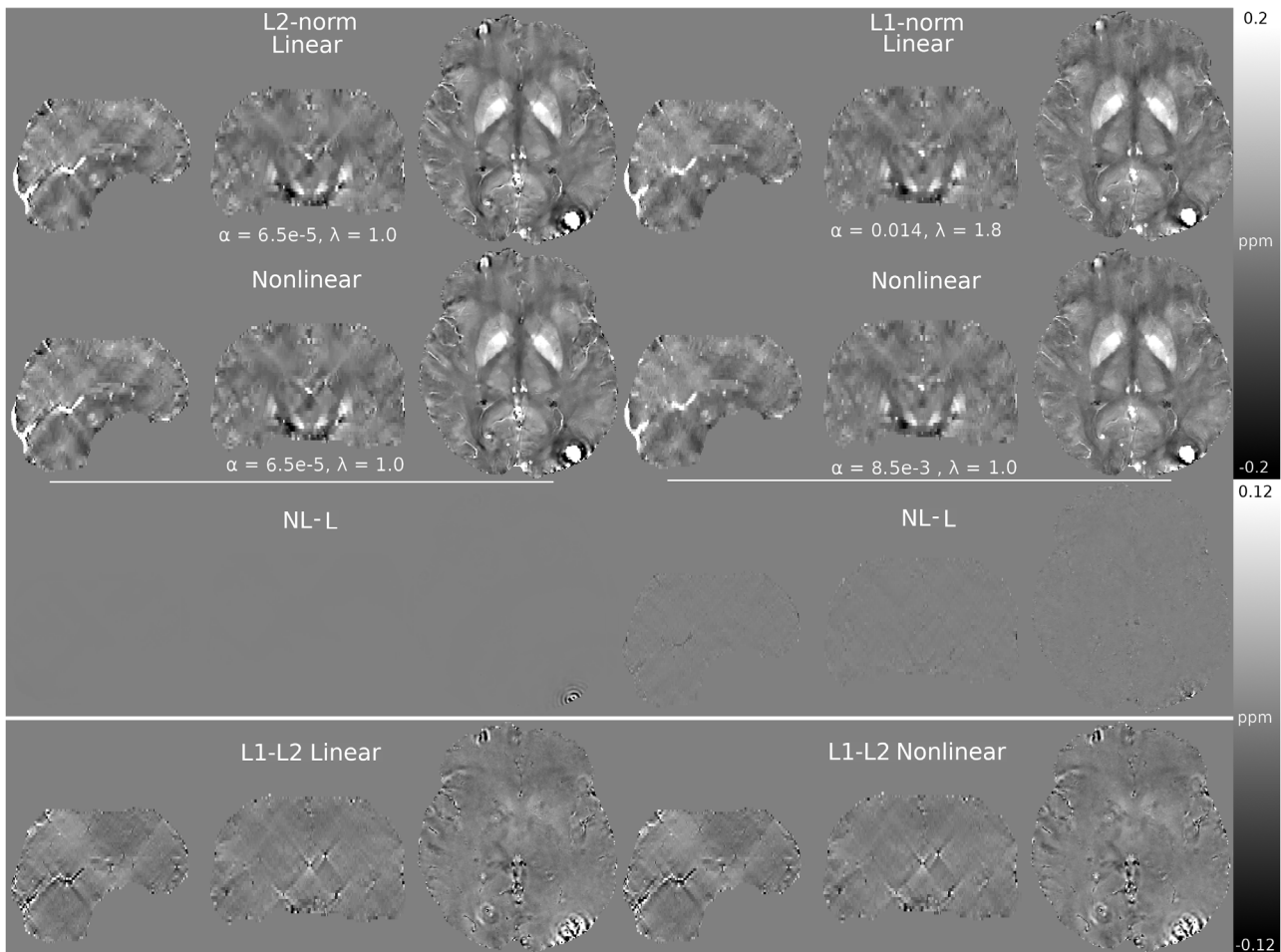
FIGURE 8 Optimal (L-curve analysis) reconstructions of the 3T in vivo dataset of a patient with a brain hemorrhage. Further fine-tuning of the λ parameter was performed visually for L1-norm methods. Difference maps between algorithms are provided below. L2-norm methods show high agreement between both results. L1-norm methods show similar streaking suppression, with different low amplitude noise management. Both streaking and ghosting artifacts are reduced in L1-norm methods, compared to L2-norm methods

artifacts. Less strong large-scale inconsistencies also propagate, often leading to so-called ghosting artifacts. To mitigate this effect, most recent developments in the field relied on either regularization terms at the cost of loss of sharpness or small-scale structures or in excluding phase inconsistencies with iterative approaches. The fundamental problem with these approaches is that Bayesian methods model the noise in the phase or the complex image domain as Gaussian. Outliers are not well handled by this model, as it enforces solutions that minimize the energy of the discrepancies. This energy is more easily dissipated along the magic cone in the frequency domain, giving birth to streaking artifacts.

In this work, we proposed a different data fidelity term based on the least absolute error model. Our experiments showed that using an L1-norm, instead of the traditional L2-norm, in the data fidelity term prevented streaking artifact propagation. In this framework, the energy of the discrepancies is not minimized, but rather the discrepancy map is forced to be (weakly) sparse. This constrains the sources of potential streaking artifacts. Synthetic experiments showed

that L1-norm based algorithms are more accurate in reconstructing results, even without using a mask to reject inconsistencies in the data fidelity term. Furthermore, optimal α is within an order of magnitude despite the masking or weighting choice in the data fidelity term, facilitating parameter optimization. L1-norm algorithms also better handled low-signal data in cortical areas, preventing noise amplification or over-regularization of the structures. This is particularly noticeable for vessels, as shown in our analytic experiments. Both L2-norm and L1-norm methods can reconstruct susceptibilities from the local phase of in vivo data with different acquisition parameters (voxel size, TE, etc), although optimal results require fine-tuning α through heuristic methods such as the L-curve analysis or the analysis of Fourier coefficients of the reconstructions.[30] Normalization of the local phase to a specific dynamic range in radians (ie, using a predefined ppm-to-radians scaling factor based on a specific TE and field strength) enables to use a more generic set of parameters,[49] but this is largely constrained by SNR and the presence of phase discrepancies. Medium and large-scale background

field residuals should be carefully avoided, as they may be reconstructed as susceptibility sources inside the ROI, instead of being considered as phase discrepancies.

A drawback of L1-norm based methods is that they are associated with a median filter[31-33] instead of a mean filter. Median filters produce overall lower SNR than mean filters when the underlying noise distribution is Gaussian. While this noisier texture is visible in the linear L1-norm results, the nonlinear L1-norm solver mostly prevented this because it works in the complex image domain, with noise amplification only apparent in very low SNR areas such as the cerebrospinal fluid (CSF). Another minor drawback of L1-norm methods is the need for fine-tuning λ, the multiplicative factor of the data-fidelity weight. This parameter controls the thresholding process in the proximal operation. Increasing its value rejects fewer outlier voxels, allowing more streaks to be generated. Too large weighting values may also lead to divergence in the nonlinear solver (typically, $\lambda > \pi/(TE\gamma B_0)$ or $\lambda > 2$). If this value is decreased too much, large contrast features such as vessels tended to be underestimated, although the solver is more stable. This may be further exploited to automate the search of the optimal weights in an incremental way (similar to the a multiscale representation[49]). In vivo, optimizing λ could be done after optimizing α, by visually ensuring that no relevant features are being suppressed. Another route to explore could be the use of preconditioners[50,51] or early-stopping[52] methods with different settings as initialization.

An example of more realistic phase inconsistency sources was presented for the SIM2SNR1 dataset, with strong intravoxel dephasing effects. This is an effect often neglected in QSM models, where a nonlinear behavior of the complex signal is generated by averaging the effects of in-homogenous sources within a voxel. The simple convolution model with the dipole kernel is unable to capture this effect, and discrepancies are created. Both L1-norm methods were more successful than their L2-norm counterparts in preventing streaks and minimizing the calcification moment's deviation. The nonlinear L1-norm solver yielded the best scores for most metrics. To contextualize our results compared to other RC2 submissions, please refer to the additional text provided in the Supporting Information. To summarize, the results shown in Figure 5 using the L1-norm methods directly compete with the overall best results. While our nonlinear L1-norm method was not developed at the time of the challenge deadlines, our linear L1-norm method achieved the best overall score in the visual assessment and was among the highest-scoring submissions.

In terms of computational cost, both L1-norm methods require slightly more memory, as an additional splitting variable was introduced. The nonlinear L1-norm method also seemed to converge slower in the internal Newton-Raphson loop, explaining an increase of 45% per iteration times. A faster outer convergence rate compensates for this.

The suggested 0.1% update rate stopping criterion was met earlier. This stopping criterion was correlated with optimal RMSE and SSIM scores, rendering it a reasonable choice for in vivo clinical settings. For oximetry purposes, a larger number of iterations may be needed for optimal representation of the susceptibility values inside the vessels. The in vivo experiments showed that L1-norm methods have a limited capacity to deal with low intensity and large-scale discrepancies, typically arising from background field and other phase residuals (from the coil combination or multi-echo fitting processes). Further extension of the functional, as proposed for the WH-TV method,[35] is a feasible alternative to account for these effects, with minor changes to the currently available source codes. As shown in Figure 7, the (nonlinear L2-norm) WH-TV QSM reconstruction was able to suppress more background residuals than subsequent L2-norm and L1-norm reconstructions using the corrected field map. Further extensions of these L1-norm methods may also be required to deal with anisotropic or microstructural effects, such as including diffusion tensor imaging information.[53,54] Also, further in vivo validation should be performed to assess the application of L1-norm methods to reconstruct cortical areas or for full field of view (FOV) (whole-head) reconstructions.

7 | CONCLUSIONS

L1-norm data fidelity term methods provided a new and efficient way to prevent streaking artifacts in QSM reconstructions by improved handling of phase outliers. The least absolute error optimization usedprevents energy spilling along the frequency coefficients in the magic cone, allowing a better reconstruction of structures in cortical and low-signal regions. Reconstruction without tissue masking also becomes feasible. L1-norm methods improved the scores of their L2-norm counterparts, with the nonlinear L1-norm method scoring the best overall numerical results. Whereas the linear L1-norm method produces results with lower SNR at homogeneous regions, results with the nonlinear version are smooth, with proper representation of high-frequency structures. This framework may be extended in the future to allow QSM studies outside of the brain or in a single-step framework.[15,35,50,55-57] Clinically, the L1 approach may enable novel challenging applications, such as assessing brain hemorrhages and cortical layers.

ORCID
Carlos Milovic https://orcid.org/0000-0002-1196-6703
Christian Langkammer https://orcid.org/0000-0002-7097-9707
Kristian Bredies https://orcid.org/0000-0001-7140-043X

REFERENCES
1. Haacke EM, Brown R, Thompson M, Venkatesan R. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*, 1st edn. Wiley-Liss; 1999.
2. Salomir R, De Senneville BD, Moonen CTW. A fast calculation method for magnetic field inhomogeneity due to an arbitrary distribution of bulk susceptibility. *Concepts Magn Reson*. 2003;19B:26-34.
3. Marques JP, Bowtell R. Application of a Fourier-based method for rapid calculation of field inhomogeneity due to spatial variation of magnetic susceptibility. *Concepts Magn Reson Part B Magn Reson Eng*. 2005;25:65-78.
4. Li L, Leigh JS. Quantifying arbitrary magnetic susceptibility distributions with MR. *Magn Reson Med*. 2004;51:1077-1082.
5. Morgan J, Irarrazaval P. Efficient solving for arbitrary susceptibility distributions using residual difference fields. *Proc 15th Annual Meeting of ISMRM (Berlin, Germany)*. 2007;35.
6. Shmueli K, de Zwart JA, van Gelderen P, Li T-Q, Dodd SJ, Duyn JH. Magnetic susceptibility mapping of brain tissue in vivo using MRI phase data. *Magn Reson Med*. 2009;62:1510-1522.
7. Wharton S, Schäfer A, Bowtell R. Susceptibility mapping in the human brain using threshold-based k-space division. *Magn Reson Med*. 2010;63:1292-1304.
8. de Rochefort L, Liu T, Kressler B, et al. Quantitative susceptibility map reconstruction from MR phase data using Bayesian regularization: validation and application to brain imaging. *Magn Reson Med*. 2010;63:194-206.
9. Kressler B, de Rochefort L, Liu T, Spincemaille P, Jiang Q, Wang Y. Nonlinear regularization for per voxel estimation of magnetic susceptibility distributions from MRI field maps. *IEEE Trans Med Imaging*. 2010;29:273-281.
10. Liu J, Liu T, de Rochefort L, et al. Morphology enabled dipole inversion for quantitative susceptibility mapping using structural consistency between the magnitude image and the susceptibility map. *Neuroimage*. 2012;59:2560-2568.
11. Bilgic B, Fan AP, Polimeni JR, et al. Fast quantitative susceptibility mapping with L1-regularization and automatic parameter selection. *Magn Reson Med*. 2014;72:1444-1459.
12. Bredies K, Kunisch K, Pock T. Total generalized variation. *SIAM J Imaging Sci*. 2010;3:492-526.
13. Knoll F, Bredies K, Pock T, Stollberger R. Second order total generalized variation (TGV) for MRI. *Magn Reson Med*. 2011;65:480-491.
14. Yanez F, Fan A, Bilgic B, Milovic C, Adalsteinsson E, Irarrazaval P. Quantitative susceptibility map reconstruction via a total generalized variation regularization. *Int Workshop Pattern Recognit Neuroimaging*. 2013;2013:203-206.
15. Langkammer C, Bredies K, Poser B, et al. Fast quantitative susceptibility mapping using 3D EPI and total generalized variation. *Neuroimage*. 2015;111:622-630.
16. Bilgic B, Chatnuntawech I, Langkammer C, Setsompop K. *Sparse Methods for Quantitative Susceptibility Mapping*. Wavelets and Sparsity XVI, SPIE; 2015. https://doi.org/10.1117/12.2188535
17. Wu B, Li W, Guidon A, Liu C. Whole brain susceptibility mapping using compressed sensing. *Magn Reson Med*. 2012;67:137-147.
18. Ahn HS, Park SH, Ye JC. Quantitative susceptibility map reconstruction using annihilating filter-based low-rank Hankel matrix approach. *Magn Reson Med*. 2020;83:858-871.
19. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. *Magn Reson Med*. 1995;34:910-914.
20. Liu T, Wisnieff C, Lou M, Chen W, Spincemaille P, Wang Y. Nonlinear formulation of the magnetic field to source relationship for robust quantitative susceptibility mapping. *Magn Reson Med*. 2013;69:467-476.
21. Wang S, Liu T, Chen W, et al. Noise effects in various quantitative susceptibility mapping methods. *IEEE Trans Biomed Eng*. 2013;60:3441-3448.
22. Milovic C, Bilgic B, Zhao B, Acosta-Cabronero J, Tejos C. Fast nonlinear susceptibility inversion with variational regularization. *Magn Reson Med*. 2018;80:814-821.
23. Zhao B, Setsompop K, Ye H, Cauley S, Wald LL. Maximum likelihood reconstruction for magnetic resonance fingerprinting. *IEEE Trans Med Imaging*. 2016;35:1812-1823.
24. Goldstein T, Osher S. The split Bregman method for L1-regularized problems. *SIAM J Imaging Sci*. 2009;2:323-343.
25. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found Trends Mach Learn*. 2010;3:1-122.
26. Kee Y, Liu Z, Zhou L, et al. Quantitative susceptibility mapping (QSM) algorithms: mathematical rationale and computational implementations. *IEEE Trans Biomed Eng*. 2017;64:2531-2545.
27. Tang J, Liu S, Neelavalli J, Cheng YCN, Buch S, Haacke EM. Improving susceptibility mapping using a threshold-based K-space/image domain iterative reconstruction approach. *Magn Reson Med*. 2013;69:1396-1407.
28. Li W, Wang N, Yu F, et al. A method for estimating and removing streaking artifacts in quantitative susceptibility mapping. *Neuroimage*. 2015;108:111-122.
29. Wei H, Dibb R, Zhou Y, et al. Streaking artifact reduction for quantitative susceptibility mapping of sources with large dynamic range. *NMR Biomed*. 2015;28:1294-1303.
30. Milovic C, Prieto C, Bilgic B, et al. Comparison of parameter optimization methods for quantitative susceptibility mapping. *Magn Reson Med*. 2020;85:480-494.
31. Arce GR. *Nonlinear Signal Processing: A Statistical Approach*. Wiley; 2005.
32. Ray N. Median filter with absolute value norm spatial regularization. Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, 2011:1437-1440.
33. Yang Q, Tan K, Ahuja N. Real-time O(1) bilateral filtering. Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, 2009;557-564.

34. Bredies K, Vicente D. A perfect reconstruction property for PDE-constrained total-variation minimization with application in Quantitative Susceptibility Mapping. *ESAIM Control Optim Calc Var*. 2019;25:83.

35. Milovic C, Bilgic B, Zhao B, Langkammer C, Tejos C, Acosta-Cabronero J. Weak-harmonic regularization for quantitative susceptibility mapping. *Magn Reson Med*. 2019;81:1399-1411.

36. Ramani S, Fessler JA. Parallel MR image reconstruction using augmented Lagrangian methods. *IEEE Trans Med Imaging*. 2011;30:694-706.

37. Liu T, Spincemaille P, De Rochefort L, Kressler B, Wang Y. Calculation of susceptibility through multiple orientation sampling (COSMOS): a method for conditioning the inverse problem from measured magnetic field map to susceptibility source image in MRI. *Magn Reson Med*. 2009;61:196-204.

38. Langkammer C, Schweser F, Shmueli K, et al. Quantitative susceptibility mapping: report from the 2016 reconstruction challenge. *Magn Reson Med*. 2018;79:1661-1673.

39. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process*. 2004;13:600-612.

40. Wang Z, Bovik AC. Mean squared error: love it or leave it? A new look at signal fidelity measures. *IEEE Signal Process Mag*. 2009;26:98-117.

41. Horé A, Ziou D. Image quality metrics: PSNR vs. SSIM. In: Proceedings International Conference on Pattern Recognition, IEEE. 2010;2366-2369.

42. Marques JP, Bilgic B, Meineke J, et al. *Towards QSM Challenge 2.0: Creation and Evaluation of a Realistic Magnetic Susceptibility Phantom*. Proc. 27th Annual Meeting of the ISMRM, Montreal, Canada: 2019:p1122.

43. Marques JP, Meineke J, Milovic C, et al. QSM reconstruction challenge 2.0: a realistic in silico head Phantom for MRI data simulation and evaluation of susceptibility mapping procedures. *Magn Reson Med*. 2021;86:526-542.

44. QSM Reconstruction Challenge Committee, Meineke J, Schweser F, et al. QSM challenge 2.0: design and report of results. *Magn Reson Med*. 2021;86:1241-1255.

45. Liu T, Khalidov I, de Rochefort L, et al. A novel background field removal method for MRI using projection onto dipole fields (PDF). *NMR Biomed*. 2011;24:1129-1136.

46. Zhou D, Liu T, Spincemaille P, Wang Y. Background field removal by solving the Laplacian boundary value problem. *NMR Biomed*. 2014;27:312-319.

47. Milovic C, Tejos C, Acosta-Cabronero J, et al. The 2016 QSM challenge: lessons learned and considerations for a future challenge design. *Magn Reson Med*. 2020;84:1624-1637.

48. Liu C. Susceptibility tensor imaging. *Magn Reson Med*. 2010;63:1471-1477.

49. Acosta-Cabronero J, Milovic C, Mattern H, Tejos C, Speck O, Callaghan MF. A robust multi-scale approach to quantitative susceptibility mapping. *Neuroimage*. 2018;183:7-24.

50. Liu Z, Kee Y, Zhou D, Wang Y, Spincemaille P. Preconditioned total field inversion (TFI) method for quantitative susceptibility mapping. *Magn Reson Med*. 2017;78:303-315.

51. Liu Z, Wen Y, Spincemaille P, et al. Automated adaptive preconditioner for quantitative susceptibility mapping. *Magn Reson Med*. 2020;83:271-285.

52. Polak D, Chatnuntawech I, Yoon J, et al. Nonlinear dipole inversion (NDI) enables robust quantitative susceptibility mapping (QSM). *NMR Biomed*. 2020;33:e4271. https://doi.org/10.1002/nbm.4271

53. Li X, van Zijl PC. Mean magnetic susceptibility regularized susceptibility tensor imaging (MMSR-STI) for estimating orientations of white matter fibers in human brain. *Magn Reson Med*. 2014;72:610-619. https://doi.org/10.1002/mrm.25322

54. Bao L, Xiong C, Wei W, Chen Z, van Zijl PCM, Li X. Diffusion-regularized susceptibility tensor imaging (DRSTI) of tissue microstructures in the human brain. *Med Image Anal*. 2021;67:101827. https://doi.org/10.1016/j.media.2020.101827

55. Chatnuntawech I, McDaniel P, Cauley SF, et al. Single-step quantitative susceptibility mapping with variational penalties. *NMR Biomed*. 2017;30:e3570. https://doi.org/10.1002/nbm.3570

56. Liu T, Zhou D, Spincemaille P, Yi W. Differential approach to quantitative susceptibility mapping without background field removal. In: Proceedings of the 22nd Annual Meeting of ISMRM, Milan, Italy: 2014:p597.

57. Sun H, Ma Y, MacDonald ME, Pike GB. Whole head quantitative susceptibility mapping using a least-norm direct dipole inversion method. *Neuroimage*. 2018;179:166-175.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

Supplementary Material

FIGURE S1 Absolute error maps for all the algorithms and methods compared in Figure 1. Note the noise amplification effect near the superior cortical area for the "ROI mask" and "Magnitude weight" methods. "Magnitude weight" produced some loss of sharpness and details in the veins at the same region (yellow arrow). Both effects are less prominent in L1-norm methods. Red arrows point out extreme errors in the boundary of the region of interest in L2-norm methods when the data fidelity term is not masked

FIGURE S2 Different orientations of the best reconstructions, for each method, in the COSMOS-based forward simulation experiment. No degradations are included (ground-state simulation)

FIGURE S3 Top: Comparisons made between results achieved at 2500 iterations and a 0.1% update as stopping criterion (see Figure 2 for the respective number of iterations, typically in the 100-300 range). Bottom: Error (difference) maps between the results of 2500 iterations and the COSMOS ground-truth. Increasing the number of iterations reduced errors in regions with higher absolute susceptibility values, such as the veins

FIGURE S4 Evolution of the RMSE (a), SSIM (b) up to a 0.1% update, and algorithm convergence up to 2500 iterations (right), in the COSMOS-based forward simulation experiment that included the 2 unmasked spheres. Phase inconsistencies with the dipole model (due to extreme noise) impedes agreement between the RMSE and SSIM metrics. A larger number of iterations tends to amplify the streaking artifacts

FIGURE S5 Reconstructions of the COSMOS-based forward simulations that included 5 strong single-voxel susceptibility sources (shown with red circles for the first algorithm). All

algorithms performed similarly, without relevant errors

FIGURE S6 Reconstructions of the RC2 SIM2SNR1 dataset using a stopping criterion of a 0.1% update ratio. A magnified view around the calcification is provided for each algorithm. Optimal reconstructions (Figure 5) showed weaker streaking originated from the calcification (red arrow)

FIGURE S7 Optimal (L-curve analysis) reconstructions of the RC1 (single orientation) dataset, using a 0.1% update as stopping criterion. Difference maps between algorithms are also provided. In all cases, the data-fidelity weight was set by the "Magnitude weight" method, with $\lambda = 1$. Differences between both L2-norm methods are minor. L1-norms methods showed larger errors in terms of overall texture (linear L1-norm seems noisier) and in the veins (such as the Galen vein, highlighted with red arrows). Nevertheless, L1-norm methods show increased contrast and details for small veins, as shown with green arrows for the nonlinear methods comparison. On top, a multi-orientation reconstruction based on a L2-norm closed-form projection of the anisotropic χ_{13} and χ_{23} susceptibility tensorial components onto the χ_{33} isotropic tensor component as described in Milovic et al[47]

TABLE S1 SSIM, RMSE, and mean absolute error (MAE) metric scores for reconstructions using noisy phase data simulated using a COSMOS-based susceptibility ground truth. RMSE and MAE scores are given in percentages, normalized by the norm of the ground truth. The weight of the data fidelity term was set according to: (a) "No mask" (full FOV included), (b) "ROI mask" (signal voids rejected), and (c) "Magnitude weight" (spatially variable weight proportional to the magnitude data). All methods were stopped when a 0.1% update was achieved

TABLE S2 SSIM and RMSE scores for the reconstructions of the COSMOS-based forward simulation, changing the number of iterations (stopped when a 0.1% update is reached, when the best RMSE score is achieved, or at a total of 2500 iterations). Linear algorithms used a ROI mask, whereas nonlinear algorithms used the magnitude weighting as data-fidelity local weight. Average times per iteration are also provided. The effective number of iterations for the 0.1% update and the best reconstruction results are presented in Figure 2