Published in final edited form as:
Magn Reson Med. 2014 November ; 72(5): 1444–1459. doi:10.1002/mrm.25029.

Fast Quantitative Susceptibility Mapping with L1-Regularization and Automatic Parameter Selection

Berkin Bilgic[1], **Audrey P. Fan**[1,2], **Jonathan R. Polimeni**[1,3], **Stephen F. Cauley**[1], **Marta Bianciardi**[1,3], **Elfar Adalsteinsson**[1,2,4], **Lawrence L. Wald**[1,3,4], and **Kawin Setsompop**[1,3]

[1]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA

[2]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, MA, USA

[3]Department of Radiology, Harvard Medical School, Boston, Massachusetts, USA

[4]Harvard-MIT Health Sciences and Technology, Cambridge, MA, USA

Abstract

Purpose—To enable fast reconstruction of quantitative susceptibility maps with Total Variation penalty and automatic regularization parameter selection.

Methods—ℓ_1-regularized susceptibility mapping is accelerated by variable-splitting, which allows closed-form evaluation of each iteration of the algorithm by soft thresholding and FFTs. This fast algorithm also renders automatic regularization parameter estimation practical. A weighting mask derived from the magnitude signal can be incorporated to allow edge-aware regularization.

Results—Compared to the nonlinear Conjugate Gradient (CG) solver, the proposed method offers 20× speed-up in reconstruction time. A complete pipeline including Laplacian phase unwrapping, background phase removal with SHARP filtering and ℓ_1-regularized dipole inversion at 0.6 mm isotropic resolution is completed in 1.2 minutes using Matlab on a standard workstation compared to 22 minutes using the Conjugate Gradient solver. This fast reconstruction allows estimation of regularization parameters with the L-curve method in 13 minutes, which would have taken 4 hours with the CG algorithm. Proposed method also permits magnitude-weighted regularization, which prevents smoothing across edges identified on the magnitude signal. This more complicated optimization problem is solved 5× faster than the nonlinear CG approach. Utility of the proposed method is also demonstrated in functional BOLD susceptibility mapping, where processing of the massive time-series dataset would otherwise be prohibitive with the CG solver.

Conclusion—Online reconstruction of regularized susceptibility maps may become feasible with the proposed dipole inversion.

Corresponding Author: Berkin Bilgic, Building 75, Room 2.102, 13th Street, Charlestown, MA, 02129, USA, berkin@nmr.mgh.harvard.edu.

Keywords

Quantitative Susceptibility Mapping; Regularization; Total Variation; L-curve

Introduction

Quantitative Susceptibility Mapping (QSM) aims to estimate the underlying magnetic susceptibility of the tissues that give rise to subtle changes in the magnetic field. Mapping this property allows quantification of tissue iron concentration (1,2), vessel oxygen saturation (3–6), and has found recent use in the investigation of neurodegenerative diseases such as multiple sclerosis (7) and characterization of brain lesions (8).

Tissue susceptibility $\boldsymbol{\chi}$ is related to the measured field map $\boldsymbol{\varphi}$ via the relation $\mathbf{DF}\boldsymbol{\chi} = \mathbf{F}\boldsymbol{\varphi}$, where $\mathbf{F}$ is the discrete Fourier Transform operator and $\mathrm{D}=1/3-k_z^2/k^2$ is the susceptibility kernel in k-space (9). This kernel effectively undersamples the frequency content of the field map on the conical surface $3k_z^2=k^2$, which makes the inversion of the relation ill-conditioned. Existing approaches improve conditioning either by imposing prior knowledge on the reconstructed susceptibility map (10–12), or by collecting additional data by changing the orientation of the sample with respect to the main field (13,14). Compared to the regularized reconstruction techniques that employ a signal prior, multi-orientation methods yield susceptibility maps with superior quality, however at the cost of increased scan time and patient discomfort (15).

Single-orientation, regularized QSM methods often impose sparsity or smoothness constraints on the spatial gradients of the reconstructed susceptibility map, which can be formulated by penalizing the ℓ_1- or ℓ_2-norm of the gradients in three dimensions (10–12,16). The two types of regularized reconstruction methods can be expressed as an unconstrained convex optimization problem, minimizing either

$$\ell_1-\text{regularized:}\frac{1}{2}\|\mathbf{F}^{-1}\mathbf{DF}\boldsymbol{\chi}-\boldsymbol{\phi}\|_2^2+\alpha\cdot\|\mathbf{WG}\boldsymbol{\chi}\|_1 \quad (1)$$

$$\ell_2-\text{regularized:}\|\mathbf{F}^{-1}\mathbf{DF}\boldsymbol{\chi}-\boldsymbol{\phi}\|_2^2+\beta\cdot\|\mathbf{WG}\boldsymbol{\chi}\|_2^2 \quad (2)$$

where α and β are regularization parameters and $\mathbf{G} = [\mathbf{G}_x; \mathbf{G}_y; \mathbf{G}_z]$ is the gradient operator in three spatial dimensions. The diagonal weighting term $\mathbf{W}$ can either be taken to be the identity matrix $\mathbf{I}$, or a binary mask $\mathbf{W} = [\mathbf{W}_x; \mathbf{W}_y; \mathbf{W}_z]$ derived from the magnitude signal can be employed to prevent penalizing certain gradient features (10–12). These two optimization problems are conventionally solved with the iterative nonlinear Conjugate Gradient algorithm. Reconstruction times reported in the literature range between 20 minutes (11,12) to 2–3 hours (17), which may be a limiting factor for the online reconstruction of susceptibility maps required for routine use or in clinical settings. Moreover, the parameters α and β need to be determined for optimal regularization, which is usually addressed by performing multiple reconstructions while sweeping a range of

parameters to trace the L-curve (18) or finding the operating point that satisfies the discrepancy principle (19). As such, identification of suitable amount of regularization would further increase the computation time.

We recently introduced a closed-form solution to the ℓ_2-regularized QSM problem in Eq.2 without magnitude weighting (**W = I**) (20). This method requires only two Fast Fourier Transform (FFT) operations and takes less than a second to compute for a high-resolution phase data. In contrast, a closed-form solution to the ℓ_1-constrained problem in Eq.1 does not exist, which forces the existing algorithms to operate iteratively. Previously, ℓ_1-regularized reconstruction was shown to be superior to ℓ_2-penalty both in image quality and quantification accuracy (12). The ℓ_1-regularized results presented herein also show substantially reduced reconstruction error on numerical simulation, and better estimation of undersampled content near the magic angle on in vivo data. In this article, we propose a fast ℓ_1-regularized QSM algorithm that works iteratively, however each iteration is computed efficiently in closed-form. We also extend these fast ℓ_1- and ℓ_2-regularized solvers to incorporate magnitude prior.

This work employs an efficient variable-splitting algorithm (21) to solve Eq.1 without magnitude weighting, and reports 20× speed up in reconstruction time compared to the nonlinear Conjugate Gradient solver (12,16). By introducing an auxiliary variable that replaces the image gradient, each iteration of the proposed algorithm is computed in closed-form, requiring only Fast Fourier Transforms (FFTs) and soft thresholding operations. With the proposed formulation, reconstruction for high-resolution *in vivo* susceptibility mapping at 0.6 mm isotropic voxel size takes 1 minute (using Matlab on a standard workstation). Combined with state-of-the-art phase unwrapping and background phase removal methods (14,22), this comprises a fast reconstruction pipeline that might facilitate clinical application of QSM. In the Discussion section, we outline the differences between the presented methods and a similar variable-splitting algorithm for QSM previously proposed in the elegant contribution by Chen et al. (23).

With the addition of magnitude prior, solution of both ℓ_1- and ℓ_2-regularized problems become more involved, since the related linear systems that need to be inverted are no longer diagonal. By employing a preconditioner that facilitates the inversion of these systems, we obtain a rapid iterative solver that leads to substantial computational savings. Relative to the nonlinear CG method, we demonstrate 5 and 15× speed-up for the magnitude-weighted ℓ_1- and ℓ_2-regularization, respectively.

Specific contributions of this work are,

i. **Fast susceptibility mapping with ℓ_1- and ℓ_2-regularization:** proposed ℓ_1-regularized algorithm is 20× faster than conventional Conjugate Gradient solver. We also introduce ℓ_1- and ℓ_2-based solvers with magnitude prior that allow edge-aware regularization, while achieving 5 and 15× speed-up over the CG approach.

ii. **Automatic regularization parameter selection:** using the L-curve heuristic, the parameter that yields the maximum curvature on the curve is selected as the optimal operating point. Sweeping the L-curve becomes practical with the proposed

fast reconstruction, whereas parameter selection with Conjugate Gradient solver would take several hours of computation.

iii. **Functional BOLD susceptibility mapping** (6,24,25): combination of fast phase unwrapping, background phase removal, regularization parameter estimation and susceptibility inversion makes reconstruction of large time-series datasets feasible, thus making investigation of functional BOLD QSM practical.

iv. **Reproducible research:** Matlab code for the complete reconstruction pipeline (Laplacian unwrapping, SHARP filtering, ℓ_2- and ℓ_1- regularized fast susceptibility mapping with magnitude weighting and parameter estimation) is included as supplementary material and made available online on the author's webpage at: http://web.mit.edu/berkin/www/Fast_QSM_Magnitude_Toolbox.zip

A copy of the source code is also deposited into a git archive at the following address: https://sourceforge.net/projects/fastqsm/

Theory

i. Proposed ℓ_2-Regularized QSM Algorithms

The global optimizer of the convex optimization problem in Eq.2 can be obtained by taking the gradient and setting it to zero, i.e.,

$$(\mathbf{F}^{-1}\mathbf{D}^2\mathbf{F}+\beta\cdot\mathbf{G}^T\mathbf{W}^2\mathbf{G})\boldsymbol{\chi}=\mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\phi} \quad (3)$$

The solution of this linear system is facilitated by evaluating the spatial gradient via multiplication in k-space rather than using convolution in image space. This can be accomplished by expressing the image gradient along the x-axis as $\mathbf{G}_x = \mathbf{F}^{-1}\mathbf{E}_x\mathbf{F}$, where $\mathbf{E}_x$ is a diagonal matrix that provides the k-space representation of the image-domain differencing operator $\delta_x - \delta_{x-1}$. $\mathbf{E}_x$ can be constructed by starting from the vector $\mathbf{v}_x$ with entries $\mathbf{v}_x(i)=1-exp(-2\pi\sqrt{-1}\cdot i/N_x)$, where $i = 0, \ldots, N_x - 1$ and N_x is the matrix size along x. When $\mathbf{v}_x$ is replicated along k_y and k_z dimensions in k-space to generate an image volume of size $N_x \times N_y \times N_z$, the matrix $\mathbf{E}_x$ can be obtained by assigning this volume to diagonal entries of $\mathbf{E}_x$. The operators $\mathbf{E}_y$ and $\mathbf{E}_z$ are similarly defined. Eq.3 then becomes

$$\mathbf{F}^{-1}(\mathbf{D}^2+\beta\cdot\mathbf{E}^H\mathbf{F}\mathbf{W}^2\mathbf{F}^{-1}\mathbf{E})\mathbf{F}\boldsymbol{\chi}=\mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\phi} \quad (4)$$

where $\mathbf{E} = [\mathbf{E}_x; \mathbf{E}_y; \mathbf{E}_z]$. We recently proposed (20) a fast solution to Eq.4 when no magnitude weighting is used ($\mathbf{W} = \mathbf{I}$). In this special case, further simplification of Eq.4 yields

$$\boldsymbol{\chi}=\mathbf{F}^{-1}\cdot(\mathbf{D}^2+\beta\cdot(\mathbf{E}_x^2+\mathbf{E}_y^2+\mathbf{E}_z^2))^{-1}\cdot\mathbf{D}\mathbf{F}\boldsymbol{\phi} \quad (5)$$

Note that the matrix inversion now involves only diagonal matrices, hence it is straightforward to evaluate. The total cost of ℓ_2-regularized QSM is then two FFT operations

and summation and multiplication of diagonal matrices. These usually take less than a second for three-dimensional high-resolution image volume.

With the inclusion of magnitude weighting **W**, the system that needs to be solved becomes,

$$(\mathbf{D}^2+\beta\cdot\mathbf{E}^H\mathbf{F}\mathbf{W}^2\mathbf{F}^{-1}\mathbf{E})\mathbf{F}\chi=\mathbf{D}\mathbf{F}\phi \quad (6)$$

Because this system is no longer diagonal, its inversion is more involved. The linear conjugate gradient algorithm is a popular technique used for solution of such symmetric, positive definite systems. The convergence speed of this method is determined by the condition number of the matrix $\mathbf{A} = (\mathbf{D}^2 + \beta\cdot\mathbf{E}^H\mathbf{F}\mathbf{W}^2\mathbf{F}^{-1}\mathbf{E})$. The conditioning can be improved if the matrix **A** can be brought closer to being identity **I**. We propose to use the closed-form solution in Eq.5 as preconditioner and solve the modified system,

$$(\mathbf{D}^2+\beta\cdot\mathbf{E}^2)^{-1}\cdot\{\mathbf{A}\cdot\mathbf{F}\chi-\mathbf{D}\mathbf{F}\phi\}=0 \quad (7)$$

Since $(\mathbf{D}^2 + \beta\cdot\mathbf{E}^2)$ is diagonal, it is straightforward to compute its inverse and convenient to use it as a preconditioner. When the gradients of the magnitude image are thresholded so that the strongest 30 % voxels within the brain mask are considered to be edges as suggested in (26), the resulting **W** matrix is equal to the identity **I** except for ~5 % of its entries. This makes the approximation $(\mathbf{D}^2 + \beta\cdot\mathbf{E}^2)^{-1} \approx \mathbf{A}^{-1}$ valid, and renders the preconditioner useful. The closed-form solution in Eq.5 can be used as an initial guess to further accelerate convergence.

Herein, the native Matlab function pcg.m is used for solving Eq.7. The major cost of each iteration is 6 FFTs required for the operator **A**.

ii. ℓ_1-Regularized QSM with Nonlinear Conjugate Gradient Solver

Nonlinear Conjugate Gradient (CG) is a popular method that is especially suitable for solution of large-scale optimization problems. It requires evaluation of only the first-order derivatives and has better convergence properties than simple gradient descent techniques. The algorithm can be summarized as follows:

Conjugate Gradient Algorithm for ℓ_1 – Regularized QSM

precomputation: $\mathbf{DF}\chi$, $\mathbf{DF}(\Delta\chi)$, $\mathbf{WG}\chi$, $\mathbf{WG}(\Delta\chi)$

line search for stepsize τ: $\operatorname{argmin}_\tau 0.5\cdot\|\mathbf{DF}(\chi+\tau\cdot\Delta\chi)-\mathbf{F}\phi\|_2^2+\alpha\|\mathbf{WG}(\chi+\tau\cdot\Delta\chi)\|_1$

χ *update*: $\chi = \chi + \tau\cdot\Delta\chi$

gradient update: $g_{\text{new}} = \mathbf{F}^{-1}\mathbf{D}(\mathbf{DF}\chi - \mathbf{F}\phi) + \alpha\cdot\mathbf{G}^T\mathbf{W}^T\operatorname{sign}(\mathbf{WG}\chi)$

parameter update: $b = \|g_{\text{new}}\|_2/\|g_{\text{prev}}\|_2$
$g_{\text{prev}} = g_{\text{new}}$

$\Delta\chi$ *update*: $\Delta\chi = b$ middot; $\Delta\chi - g_{\text{new}}$

Note that the data consistency term for the CG algorithm is taken to be $\|\mathbf{DF}\chi-\mathbf{F}\phi\|_2^2$ instead of $\|\mathbf{F}^{-1}\mathbf{DF}\chi-\phi\|_2^2$ as in Eq.1. Due to Parseval's relation, the two terms are equivalent, however using this alternative formulation avoids four redundant FFT operations per iteration. The algorithm terminates when a user-specified convergence criterion is met. Per iteration, it requires four FFT operations and four spatial gradient evaluations (two for *precomputation* and two for *gradient update*). Since each gradient computation involves finite differences in three dimensions, total cost of the algorithm is twelve differencing operations and four FFT operations per iteration.

iii. Proposed ℓ_1-Regularized QSM Algorithms

The proposed approach relies on a fast algorithm that was initially proposed for reconstruction of undersampled k-space data in the context of compressed sensing (21). Here, this method is extended to susceptibility mapping using the following formulation,

$$min_{\chi,y}\frac{1}{2}\|\mathbf{F}^{-1}\mathbf{DF}\chi-\phi\|_2^2+\lambda\cdot\|y\|_1 \quad such\ that\ \mathbf{WG}\chi=y \quad (8)$$

This constrained objective is first relaxed to an unconstrained optimization problem, then solved with the two-phase "split-Bregman" iteration:

at iteration t,

$$(\chi_{t+1},y_{t+1})=argmin_{\chi,y}\frac{1}{2}\|\mathbf{F}^{-1}\mathbf{DF}\chi-\phi\|_2^2+\lambda\|\mathbf{y}\|_1+\frac{\mu}{2}\|y-\mathbf{WG}\chi-\eta_t\|_2^2 \quad (9)$$

$$\eta_{t+1}=\eta_t+\mathbf{WG}\chi_{t+1}-y_{t+1} \quad (10)$$

Here, y is an auxiliary variable that replaces the magnitude-weighted gradient of the susceptibility map χ. This idea of variable splitting will allow closed-form optimization via the soft thresholding operator. The variable η adds the mismatch in $\mathbf{WG}\chi=y$ back to the unconstrained problem to enforce this constraint. Eq.9 can be solved efficiently by iteratively minimizing with respect to χ and y separately:

$$\chi_{t+1}=argmin_{\chi}\frac{1}{2}\|\mathbf{F}^{-1}\mathbf{DF}\chi-\phi\|_2^2+\frac{\mu}{2}\|\mathbf{y}_t-\mathbf{WG}\chi-\eta_t\|_2^2 \quad (11)$$

$$y_{t+1}=argmin_y\,\lambda\cdot\|y\|_1+\frac{\mu}{2}\|y-\mathbf{WG}\chi_{t+1}-\eta_t\|_2^2 \quad (12)$$

The optimality condition for Eq.11 can be found by taking the gradient and setting it to zero,

$$(\mathbf{F}^{-1}\mathbf{D}^2\mathbf{F}+\mu\cdot\mathbf{G}^T\mathbf{W}^2\mathbf{G})\chi_{t+1}=\mathbf{F}^{-1}\mathbf{DF}\phi+\mu\cdot\mathbf{G}^T\mathbf{W}^T(y_t-\eta_t) \quad (13)$$

For the special case where magnitude weighting is not utilized (**W** = **I**), the solution to Eq.13 can be evaluated in closed-form to obtain the following update rule for $\boldsymbol{\chi}$:

$$\mathbf{F}\boldsymbol{\chi}_{t+1}=(\mathbf{D}^2+\mu\cdot(\mathbf{E}_x^2+\mathbf{E}_y^2+\mathbf{E}_z^2))^{-1}\cdot\left(\mathbf{D}\mathbf{F}\boldsymbol{\phi}+\mu\cdot\mathbf{E}^H\mathbf{F}(\boldsymbol{y}_t-\boldsymbol{\eta}_t)\right) \quad (14)$$

Here, $\mathbf{E}^H=[\mathbf{E}_x^H,\mathbf{E}_y^H,\mathbf{E}_z^H]$ and the variables $\boldsymbol{y}_t$ and $\boldsymbol{\eta}_t$ have components in three dimensions, $\boldsymbol{y}_t=[\boldsymbol{y}_{t,x};\boldsymbol{y}_{t,y};\boldsymbol{y}_{t,z}]$ and $\boldsymbol{\eta}_t=[\boldsymbol{\eta}_{t,x};\boldsymbol{\eta}_{t,y};\boldsymbol{\eta}_{t,z}]$. The matrix inversion involves only diagonal matrices and needs to be computed once. The cost of evaluating Eq.14 is three FFT operations (one for each spatial axis) and element-wise multiplications. The update for the susceptibility map is performed in k-space, thereby avoiding one redundant inverse FFT operation per iteration.

In the presence of magnitude weighting, Eq.13 no longer admits a closed-form solution. However, further simplification leads to a system similar to Eq.6 encountered while solving the magnitude-weighted ℓ_2-based problem,

$$(\mathbf{D}^2+\mu\cdot\mathbf{E}^H\mathbf{F}\mathbf{W}^2\mathbf{F}^{-1}\mathbf{E})\mathbf{F}\boldsymbol{\chi}_{t+1}=\mathbf{D}\mathbf{F}\boldsymbol{\phi}+\mu\cdot\mathbf{E}^H\mathbf{F}\mathbf{W}^T(\boldsymbol{y}_t-\boldsymbol{\eta}_t) \quad (15)$$

This equation can be efficiently solved with the linear conjugate gradient method by again employing the preconditioner ($\mathbf{D}^2+\mu\cdot\mathbf{E}^2$). Convergence behavior can be substantially enhanced if the solution from the previous iteration $\boldsymbol{\chi}_t$ is used as an initial guess. As the variable-splitting algorithm iterates, the difference between the successive solutions $\boldsymbol{\chi}_t$ and $\boldsymbol{\chi}_{t+1}$ decrease progressively. Using such a suitable initial guess, preconditioned conjugate gradient can solve Eq.15 within 1% tolerance in a couple of steps.

Regardless of the magnitude weighting, solution to ℓ_1-regularized least squares problem in Eq.12 is given by the element-wise soft thresholding operation,

$$\boldsymbol{y}_{t+1}=max\left(|\mathbf{W}\mathbf{G}\boldsymbol{\chi}_{t+1}+\boldsymbol{\eta}_t|-\frac{\lambda}{\mu},0\right)\cdot sign(\mathbf{W}\mathbf{G}\boldsymbol{\chi}_{t+1}+\boldsymbol{\eta}_t) \quad (16)$$

This operator is applied to all three components of $\boldsymbol{y}_{t+1}$. To compute the gradient in x direction $\mathbf{G}_x\boldsymbol{\chi}_{t+1}$, the k-space representation of the current susceptibility estimate $\mathbf{F}\boldsymbol{\chi}_{t+1}$ is utilized due to $\mathbf{G}_x\boldsymbol{\chi}_{t+1}=\mathbf{F}^{-1}\mathbf{E}_x(\mathbf{F}\boldsymbol{\chi}_{t+1})$. This way, the update rule in Eq.16 requires three inverse FFT operations (one for each spatial axis) and simple thresholding operations.

Combining the solutions to Eqs. 11, 12 with the update rule in Eq.16 yields the complete algorithm for the proposed ℓ_1-regularized reconstruction:

Proposed Algorithm for ℓ_1 – Regularized QSM

initialization: $\boldsymbol{y}_{0,x}=\boldsymbol{y}_{0,y}=\boldsymbol{y}_{0,z}=\boldsymbol{\eta}_{0,x}=\boldsymbol{\eta}_{0,y}=\boldsymbol{\eta}_{0,z}=\mathbf{0}$

$\mathbf{F}\boldsymbol{\chi}$ *update*: **Solve** $(\mathbf{D}^2+\mu\cdot\mathbf{E}^H\mathbf{F}\mathbf{W}^2\mathbf{F}^{-1}\mathbf{E})\mathbf{F}\boldsymbol{\chi}_{t+1}=\mathbf{D}\mathbf{F}\boldsymbol{\varphi}+\mu\cdot\mathbf{E}^H\mathbf{F}\mathbf{W}^T(\boldsymbol{y}_t-\boldsymbol{\eta}_t)$

$$\textit{gradient update}: \begin{cases} \mathbf{W}_\chi \mathbf{G}_\chi \boldsymbol{\chi}_{t+1} = \mathbf{W}_\chi \mathbf{F}^{-1} \mathbf{E}_\chi (\mathbf{F}\boldsymbol{\chi}_{t+1}) \\ \mathbf{W}_y \mathbf{G}_y \boldsymbol{\chi}_{t+1} = \mathbf{W}_y \mathbf{F}^{-1} \mathbf{E}_y (\mathbf{F}\boldsymbol{\chi}_{t+1}) \\ \mathbf{W}_z \mathbf{G}_z \boldsymbol{\chi}_{t+1} = \mathbf{W}_z \mathbf{F}^{-1} \mathbf{E}_z (\mathbf{F}\boldsymbol{\chi}_{t+1}) \end{cases}$$

$$\textit{soft thresholding}: \begin{cases} \boldsymbol{y}_{t+1,x} = max(|\mathbf{W}_x \mathbf{G}_x \boldsymbol{\chi}_{t+1} + \boldsymbol{\eta}_{t,x}| - \lambda/\mu, 0) \cdot sign(\mathbf{W}_x \mathbf{G}_x \boldsymbol{\chi}_{t+1} + \boldsymbol{\eta}_{t,x}) \\ \boldsymbol{y}_{t+1,y} = max(|\mathbf{W}_y \mathbf{G}_y \boldsymbol{\chi}_{t+1} + \boldsymbol{\eta}_{t,y}| - \lambda/\mu, 0) \cdot sign(\mathbf{W}_y \mathbf{G}_y \boldsymbol{\chi}_{t+1} + \boldsymbol{\eta}_{t,y}) \\ \boldsymbol{y}_{t+1,z} = max(|\mathbf{W}_z \mathbf{G}_z \boldsymbol{\chi}_{t+1} + \boldsymbol{\eta}_{t,z}| - \lambda/\mu, 0) \cdot sign(\mathbf{W}_z \mathbf{G}_z \boldsymbol{\chi}_{t+1} + \boldsymbol{\eta}_{t,z}) \end{cases}$$

$$\textit{residual update}: \begin{cases} \boldsymbol{\eta}_{t+1,x} = \boldsymbol{\eta}_{t,x} + \mathbf{W}_x \mathbf{G}_x \boldsymbol{\chi}_{t+1} - \boldsymbol{y}_{t+1,x} \\ \boldsymbol{\eta}_{t+1,y} = \boldsymbol{\eta}_{t,y} + \mathbf{W}_y \mathbf{G}_y \boldsymbol{\chi}_{t+1} - \boldsymbol{y}_{t+1,y} \\ \boldsymbol{\eta}_{t+1,z} = \boldsymbol{\eta}_{t,z} + \mathbf{W}_z \mathbf{G}_z \boldsymbol{\chi}_{t+1} - \boldsymbol{y}_{t+1,z} \end{cases}$$

Per iteration, the proposed method without magnitude weighting requires six FFT operations (three for $\mathbf{F}\boldsymbol{\chi}$ *update* due to Eq.14 and another three for *gradient update*) and simple thresholding operations. Compared to the nonlinear Conjugate Gradient method, there is no need to compute the twelve differencing operations in image space, however there are two additional FFT operations per iteration. On the other hand, one iteration of the proposed method with magnitude prior requires (6 + 6*s*) FFTs, where *s* is the number of steps it takes to solve Eq.15.

Moreover, the proposed method introduces a second regularization parameter μ that must be determined. Ref. (21) shows that this parameter does not change the solution to the optimization problem, but affects the convergence speed. Empirically, a parameter that leads to fast convergence can be identified on one subject, which might then generalize to different datasets. Alternatively, an automatic way to select this parameter will be proposed in the Methods section, which was observed to yield fast convergence speed.

Methods

The performance of the proposed method was validated on a numerical phantom with known susceptibility. In addition, two *in vivo* datasets were reconstructed: a high-resolution whole brain 3D gradient echo (GRE) acquired at 3T with 0.6 mm isotropic resolution, and a whole brain gradient echo 2D echo-planar at 7T with 1.5 mm isotropic resolution acquired as a time series. Subjects were scanned under the approval of the local Institutional Review Board. All computations were performed in Matlab using a workstation with 32 processors (AMD Opteron 6282 SE) and 128 GB memory.

1. Numerical Susceptibility Phantom

A three-compartment phantom of size 246×246×162 with known susceptibility was generated. Within each compartment, susceptibility was fixed and constant at χ_{gray} = −0.023, χ_{white} = 0.027 and χ_{csf} = −0.018 ppm (27). From this distribution $\boldsymbol{\chi}$, the field map $\boldsymbol{\varphi}$ was generated using forward dipole model $\boldsymbol{\varphi} = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\chi}$, to which random Gaussian i.i.d. noise with a peak signal to noise ratio (PSNR) of 100 was added (Fig. 1a), where PSNR was defined to be the ratio of the maximum value in the field map to the noise standard

deviation. Starting from this noisy field map, susceptibility maps were reconstructed using three different methods:

i. ℓ_2-regularized QSM with closed-form solution,

ii. ℓ_1-regularization with nonlinear Conjugate Gradient solver, and

iii. Proposed ℓ_1-regularization with variable-splitting.

The regularization parameters (β for ℓ_2-regularization, a for Conjugate Gradient and λ for the proposed method) were determined by parameter sweeping and the values that minimized the normalized root mean square error (RMSE) relative to the true $\boldsymbol{\chi}$ were selected to be the optimal setting.

At the first iteration of the proposed algorithm, the initial estimate of the susceptibility map is found by using the initial condition $\boldsymbol{y}_0 = \boldsymbol{\eta}_0 = \mathbf{0}$ in Eq.14. This yields

$$\mathbf{F}\boldsymbol{\chi}_1=(\mathbf{D}^2+\mu\cdot(\mathbf{E}_x^2+\mathbf{E}_y^2+\mathbf{E}_z^2))^{-1}\cdot\mathbf{DF}\boldsymbol{\phi} \quad (17)$$

Note that this is exactly the same expression as Eq.5, which gave the optimizer for the ℓ_2-regularized QSM formulation. Hence, it is seen that the first iteration of the proposed ℓ_1-constrained algorithm is actually the solution to the original ℓ_2-regularized problem in Eq.2. Based on this, we propose to set μ to the optimal value of β found for the closed-form ℓ_2-reconstruction. The algorithm terminates when the change in k-space of the susceptibility falls below 1%.

To hasten convergence, nonlinear Conjugate Gradient algorithm used the closed-form ℓ_2-reconstruction as initial guess, and the termination criterion was again less 1% change in susceptibility.

Ref. (21) shows that the parameter μ controlling the gradient consistency term $\mu\cdot\|\boldsymbol{y}-\mathbf{G}\boldsymbol{\chi}-\boldsymbol{\eta}_t\|_2^2$ in the variable-splitting algorithm does not affect the solution but the speed of convergence. To empirically test this property, the numerical phantom was also reconstructed with various μ settings while keeping the ℓ_1-parameter λ fixed at 10^{-5}. To enforce convergence, the algorithm was run for 300 iterations.

2. High-Resolution *In Vivo* 3D GRE

A 26 year old, female, healthy volunteer (having given informed consent) was scanned using a dual-echo gradient echo sequence with full flow compensation along all three axes (28) at a Siemens 3T system equipped with 32 receive channels. Imaging parameters were: 0.6 mm isotropic resolution, repetition time (TR) = 26 ms, echo times (TE) = 8.1 and 20.3 ms, matrix size = 384×336×224, GRAPPA acceleration factor = 2, phase partial Fourier = 75% and acquisition time = 15:42 min. Additional GRE data were collected with the same spatial coverage but at a lower resolution of 1.8×1.8×2.4 mm^3 at five different TEs (6 to 10 ms with 1 ms intervals). The radiofrequency (RF) phase offset of each channel at TE = 0 was estimated using a linear fit to the phase images corresponding to the five acquired echoes. The estimated offset per channel was then removed from the 0.6 mm high-resolution phase

data (29). Finally, magnitude-weighted, phase-valued coil combination (30) was performed and is given by

$$\psi_{combo} = \angle\left(\sum_i m_i \cdot \exp\left(\sqrt{-1}\cdot\psi_i\right)\right) \quad (18)$$

Combined phase image at TE = 8.1 ms was masked using a binary-valued region of interest (ROI) generated with FSL-BET (31) (Fig. 2a). This wrapped-phase image was unwrapped with the fast Laplacian unwrapping algorithm (22) (Fig. 2b) and the background contributions due to air-tissue interfaces were removed with SHARP filtering (14) (Fig. 2c). The truncation level for SHARP was set to 0.05 as in (14) and following (32), the kernel size was set to 9×9×9 pixels so that the filter length was about 5 mm isotropic.

The regularization parameter β for closed-form ℓ_2-regularization without magnitude prior was selected using the L-curve heuristic (18) (Fig. 3). Since the L-curve traced by the data consistency and regularization terms as β varied did not have a clear elbow point, the point with the largest curvature was selected as the optimal operating point. The curvature κ was computed as in (18),

$$\kappa = 2\frac{\ddot{\rho}\,\dot{\omega} - \ddot{\omega}\,\dot{\rho}}{\left(\dot{\rho}^2 + \dot{\omega}^2\right)^{1.5}} \quad (19)$$

Here, $\rho = \log\left(\left\|\mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\chi} - \boldsymbol{\phi}\right\|_2^2\right)$ and $\omega = \log\left(\left\|\mathbf{G}\boldsymbol{\chi}\right\|_2^2\right)$ and $\dot{\omega}$, $\dot{\rho}$, $\ddot{\omega}$ and $\ddot{\rho}$ represent first and second derivatives of ω and ρ with respect to β. Since the L-curve is only sampled at discrete points, we fit cubic splines to ρ and ω to express them as smooth functions of β, which can then be differentiated.

Proposed magnitude-weighted ℓ_2-based reconstruction with the preconditioned conjugate gradient solver used the same β parameter. The termination criterion for this iterative approach was to attain 0.1 % tolerance, where tolerance is defined as $\|\mathbf{A}\boldsymbol{x} - \boldsymbol{b}\|_2/\|\boldsymbol{b}\|_2$ regarding the solution of the system $\mathbf{A}\boldsymbol{x}=\boldsymbol{b}$. The magnitude prior was derived by thresholding the spatial gradients of the magnitude signal so that 30 % voxels in the brain mask with the strongest gradients were considered to be edges (26).

Regarding the proposed ℓ_1-based method without magnitude weighting, the smoothing parameter λ was similarly determined by maximizing the curvature of the L-curve (Fig. 4). Each point on the L-curve was reconstructed using 10 iterations of the proposed algorithm. Based on the observation that the initial iteration yields the ℓ_2-regularized reconstruction, the consistency parameter μ was set to the optimal β determined for the closed-form solution. The algorithm terminated when there was less than 1 % change in the k-space of the susceptibility map.

The dependence of convergence speed of the proposed ℓ_1-based method on gradient consistency μ parameter was also investigated. Reconstructions with parameters that are 10-

times larger and 10-times smaller than the optimal choice of μ were performed with the same 1 % change convergence criterion.

Reconstruction with ℓ_1-regularization using magnitude prior employed the same λ parameter identified from the L-curve analysis. The tolerance of the preconditioned conjugate gradient used at each iteration was set to 1 %, and the algorithm terminated when the updates between iterations was below 1 %.

The smoothing parameter α for the ℓ_1-regularized nonlinear Conjugate Gradient method was selected so that the data consistency of the final susceptibility map, $\|\mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\chi}-\boldsymbol{\phi}\|_2^2$, matched the consistency of the proposed reconstruction (Fig. 5). This strategy aimed to obtain comparable amounts of regularization with the two ℓ_1-penalized algorithms. To facilitate convergence, closed-form ℓ_2-reconstruction was used as initial guess to start the CG iterations, which terminated when there was less than 1 % change in the susceptibility map.

3. *In Vivo* 2D EPI for Functional BOLD QSM

To emulate functional BOLD QSM data acquisition, a 35 year old, male, healthy volunteer (having given informed consent) was scanned with a 2D echo-planar imaging trajectory using a Siemens 7T system with a 32-channel receive coil array (33) and a birdcage transmit coil at 1.5 mm isotropic resolution. The imaging parameters were: TR = 3660 ms, TE = 21 ms, matrix size = 128×128×76, total number of time points = 30, GRAPPA acceleration factor = 3 (reconstruction performed offline in Matlab), slice acquisition order: ascending, slice orientation: axial. Following (30), the relative phase offset of each coil image was estimated from the center 3×3×3 pixels and removed from each coil. Magnitude-weighted phase-valued coil combination was performed using Eq.18. A brain mask was generated using FSL BET (31) based on the first frame in the time series. After masking, coil-combined wrapped-phase image of each frame was processed with Laplacian unwrapping (22) and SHARP filtering (14) (Fig. 7) with truncation level = 0.05 and kernel size = 3×3×3, so that the resulting filter size was about 5 mm (32).

The resulting tissue phase series was inverted using closed-form ℓ_2-regularized QSM and the proposed ℓ_1-based reconstruction without magnitude prior. Smoothing parameters β and λ were selected with L-curve analysis using the mean tissue phase averaged over the 30 frames as the input. The gradient consistency parameter μ for the ℓ_1-based method was again set to the β value determined by L-curve. The proposed iterative algorithm terminated when there was less than 1% signal change.

To quantify the stability of phase and susceptibility time-series, time-SNR (t-SNR) and standard deviation maps were computed for the raw unwrapped phase, tissue phase after background removal, and ℓ_1- and ℓ_2-regularized susceptibility maps. The “noise” term involved in t-SNR computation was estimated by subtracting the time-average from each time point in the series. Taking the standard deviation of noise over time yielded the standard deviation maps. The “signal” term in the t-SNR metric was obtained by averaging the data over time, then taking the absolute value of this result. The ratio of the signal and standard deviation components yielded the t-SNR maps.

Results

1. Numerical Susceptibility Phantom

Starting from the noisy field map in Fig. 1a, the closed-form ℓ_2-reconstruction was completed in 0.3 seconds with 17.5 % RMSE (Figs. 1b and c). The nonlinear Conjugate Gradient method converged in 50 iterations and 258 seconds with 6.1 % error (Figs. 1d and e). The proposed variable-splitting reconstruction was completed in 10 iterations and 13 seconds with 6.7 % RMSE (Figs. 1f and g). The parameter setting that was used in these experiments was $\beta = 2.2 \cdot 10^{-4}$ (ℓ_2-recon), $\alpha = 1.5 \cdot 10^{-5}$ (CG), $\lambda = 10^{-5}$ (proposed) and these values were chosen to minimize the reconstruction error of each method.

When the proposed algorithm was run for 20 iterations (10 extra iterations), the error decreased to 6.1 % and the reconstruction time increased to 25 seconds.

To test the effect of the parameter μ in the final susceptibility solution, a wide range of μ values were swept with 300 iterations. The value of $\mu = 2.2 \cdot 10^{-4}$ used in Fig. 1f yielded 5.95 % RMSE. The same error of 5.95 % was obtained at $\mu = 2.2 \cdot 10^{-3}$ and $2.2 \cdot 10^{-2}$, while this was 6.02 % for $2.2 \cdot 10^{-5}$.

2. High-Resolution *In Vivo* 3D GRE

The proposed fast phase processing pipeline allowed unwrapping and background removal to be completed in 13 seconds for the high-resolution anatomical dataset (Fig. 2). Tracing the L-curve with the ℓ_2-based method took 42 seconds, where 15 reconstructions were made for logarithmically-spaced β values between 10^{-3} and 1 (Fig. 3 top panel). The setting that maximized the curvature of the L-curve was found to be $\beta = 3.2 \cdot 10^{-2}$, which was taken to the optimal smoothing parameter value (Fig. 3b). Two additional reconstructions with under- and over-regularization are depicted in Figs. 3a and c for comparison. Computation time for each ℓ_2-based susceptibility map was 0.9 seconds.

The smoothing parameter value corresponding to the largest curvature on the L-curve for proposed ℓ_1-regularization was found to be $\lambda = 9.2 \cdot 10^{-4}$. Total reconstruction time for 15 reconstructions with logarithmically-spaced λ values between 10^{-4} and $10^{-2.5}$ was 710 seconds (Fig. 4 top panel). Under-, optimally- and over-regularized susceptibility maps are plotted in Figs. 4a to c for comparison. Processing time for the optimally-smoothed reconstruction was 60 seconds and convergence criterion was met in 13 iterations.

The convergence speed of the ℓ_1-based method was seen to be dependent on the selection of μ parameter. The result reported in Fig. 4b was obtained with the heuristically selected value of $\mu = 3.2 \cdot 10^{-2}$. If the value of μ were instead chosen to be 10 times larger ($3.2 \cdot 10^{-1}$), the 1 % convergence criterion would be met in 20 iterations and 91 seconds. Using a parameter value 10 times smaller than the optimal ($3.2 \cdot 10^{-3}$), it would take 25 iterations and 114 seconds to converge. These are both slower than the optimal setting that required 13 iterations and 60 seconds.

The nonlinear CG smoothing parameter value that yielded matching data consistency with the proposed method was found to be $\alpha = 10^{-3}$. Computation took 1350 seconds for the

nonlinear Conjugate Gradient method, corresponding to 50 iterations (Fig. 5b). If the regularization parameter α had been estimated with the L-curve technique, the parameter sweep would have taken ~4 hours with the CG algorithm.

Regularized susceptibility maps with and without magnitude prior are presented in Fig. 6. To facilitate comparison, maximum intensity projections are taken over 3 mm thick slabs. Compared to the closed-form solution in Fig. 6a that was computed in 0.9 seconds, ℓ_2-based reconstruction with magnitude prior in Fig. 6b took 88 seconds and 14 iterations. While the proposed ℓ_1-based method without magnitude weighting (Fig. 6c) converged in 60 seconds, the processing time increased to 275 seconds when magnitude prior was included (Fig. 6d). To demonstrate the effect of magnitude weighting, sagittal planes are further zoomed in.

Maximum intensity projections of tissue phase, ℓ_2- and ℓ_1-based susceptibility maps with and without magnitude prior thorough axial, coronal and sagittal planes are depicted in Fig. 7. The corresponding k-space views are plotted in Fig. 8.

3. *In Vivo* 2D EPI for Functional QSM

Obtaining the tissue phase for the 30 time frames using Laplacian unwrapping and SHARP filtering took 9 seconds in total (0.3 seconds per frame, Fig. 9). L-curve parameter value estimation was performed once on the average tissue phase, and the resulting parameter values were applied to the 30 frames. ℓ_2-based parameter estimation took 2.7 seconds, while tracing the L-curve took 44 seconds with the proposed ℓ_1-based method. The ranges of the parameter sweeps were again $\beta \in \{10^{-3}, 1\}$ and $\lambda \in \{10^{-4}, 10^{-2.5}\}$ and the optimal parameter values were found to be $\beta = 3.2 \cdot 10^{-2}$ and $\lambda = 9.2 \cdot 10^{-4}$.

Closed-form ℓ_2-reconstruction completed in a total of 2.1 seconds (0.07 seconds per frame, Fig. 10a), whereas the proposed method completed in 192 seconds (6.4 seconds per frame, Fig. 10b).

Time-series standard deviation maps for phase and susceptibility signals are depicted in Figs.10c–f. The standard deviation and time-SNR values averaged over the brain mask were $\sigma = 6 \cdot 10^{-3}$ and t-SNR=19.1 for the raw phase (Fig. 10c), $\sigma = 3.3 \cdot 10^{-3}$ and t-SNR=3.8 for the tissue phase (Fig. 10d), $\sigma = 4.4 \cdot 10^{-3}$ and t-SNR=5.4 for ℓ_2-regularized susceptibility maps (Fig. 10e), and $\sigma = 4.7 \cdot 10^{-3}$ and t-SNR=4.3 for ℓ_1-based susceptibility reconstruction (Fig. 10f).

Regularization parameters for the QSM algorithms employed in the reconstruction of the three datasets are summarized in Table 1.

Discussion

This report introduces a fast Total Variation regularized QSM algorithm that achieves 20× processing speed-up relative to conventional nonlinear Conjugate Gradient solver. Compared to the previously reported computation times (20 minutes to several hours (12,17)), the proposed algorithm drastically reduces the processing time down to 1 minute for a whole brain high-resolution dataset. This efficient formulation of ℓ_1-constrained reconstruction might therefore facilitate the clinical use of regularized susceptibility

mapping. The increased computational efficiency stems from two sources: To reach the same convergence criterion, the proposed method requires about 4 times smaller number of iterations. At each iteration, this algorithm uses 6 FFT operations whereas the nonlinear CG solver needs 4 FFTs and 12 spatial differencing operations. The combined effect of smaller number of iterations at reduced computational load leads to the observed speed-up. The convergence criterion employed for the nonlinear CG algorithm was less than 1 % change in the image-space representation of the susceptibility map, while this criterion was 1 % change in the k-space representation of the susceptibility reconstruction for the split-Bregman algorithm. More explicitly, we compute $\|\chi_{new} - \chi_{prev}\|_2/\|\chi_{new}\|_2$ for nonlinear CG, and $\|\mathbf{F}\cdot\chi_{new} - \mathbf{F}\cdot\chi_{prev}\|_2/\|\mathbf{F}\cdot\chi_{new}\|_2$ for split-Bregman reconstruction. While the two stopping criteria are seemingly different, they are in fact equivalent due to Parseval's theorem, as $\|\mathbf{F}\cdot\chi\|_2 = \|\chi\|_2$.

Combined with the previously proposed Laplacian phase unwrapping and SHARP filtering steps (14,22), the proposed algorithm comprises a fast processing pipeline that produces susceptibility maps from the wrapped phase data. As each dipole inversion can be completed rapidly, automatic estimation of the regularization parameter also becomes practical. The proposed method without magnitude prior is capable of tracing the L-curve faster than the time it would take the Conjugate Gradient solver to complete a single reconstruction. For *in vivo* whole brain susceptibility mapping at 0.6 mm isotropic resolution, the complete pipeline of phase processing, parameter estimation and regularized QSM takes under 14 minutes. As the optimal smoothing parameters would be expected to remain within a certain range across different datasets, it would be possible to further reduce the parameter estimation time by considering a smaller interval of values. For studies that involve scanning multiple subjects with similar imaging parameters, it might also be possible to determine the regularization amount on one subject, and apply this to the rest of the subjects in the study.

Variable-splitting methods are well studied and popular in signal processing community. These are also known as *split-Bregman* iterations (21) and have been deployed successfully in compressed sensing applications for MRI (34). A variable-splitting method for susceptibility mapping similar to the proposed algorithm was proposed in (23), which aims to solve,

$$min_{\chi,y,v}\frac{1}{2}\|v-\mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi-\mathbf{z}_t\|_2^2+\lambda\cdot\|y\|_1+\frac{\mu}{2}\|y-\mathbf{G}\chi-\eta_t\|_2^2+\frac{\gamma}{2}\|v-\phi\|_2^2 \quad (20)$$

Compared to our formulation in Eq.9, this introduces a new variable ν and a third regularization parameter γ. Since the tissue phase φ is the observed data, it should not be necessary to further introduce an auxiliary variable ν to penalize these observations. By omitting the term $\frac{\gamma}{2}\|v-\phi\|_2^2$, we avoid selection of the third parameter γ and simplify the reconstruction. Further, we extend the variable-splitting algorithm to admit magnitude weighting on the spatial gradients. Originally proposed in the influential MEDI papers (10–12), magnitude-weighted regularization aims to avoid smoothing across strong edges identified on the magnitude signal. Since the linear system involved in the reconstruction is no longer diagonal, we propose to use a preconditioned conjugate gradient solver for rapid

matrix inversion. Despite the increased complexity of the optimization problem, the resulting ℓ_1-regularized algorithm with magnitude prior is still 5× faster than the nonlinear CG technique.

Herein, ℓ_2-regularized QSM is also extended to admit prior information derived from the magnitude signal. While the resulting optimization problem is no longer solved in closed-form, the proposed preconditioned linear solver allows significant computational savings. To reach a tolerance level of 0.1 %, preconditioned solver requires 14 iterations, which take 1.5 minutes for the high-resolution in vivo dataset. Without the preconditioner, the same tolerance level would have been reached with 30 iterations and more than twice the processing time. At 0.1 % tolerance, the objective value of the ℓ_2-penalized minimization problem obtained with the proposed method was 381.9675. If the nonlinear CG algorithm were used, the objective value at the end of reconstruction would have been 381.9684, with a processing time of 22 minutes. Based on this, the proposed ℓ_2-based algorithm with magnitude prior is able to solve the optimization problem 15× faster compared to nonlinear CG with similar accuracy.

Existing fast QSM algorithms include Thresholded K-space Division (TKD) (35,36) and the closed-form ℓ_2-constrained method that our group recently proposed (20). Although these techniques are computationally very efficient (requiring only two FFT operations), their regularization approaches might hamper the conspicuity of high spatial frequency components, such as vessels and small iron-rich structures like the red nucleus. This point can be appreciated on Figs.7b and d, where the vessels in the MIP image for ℓ_2-penalized reconstruction are visibly dimmer than their ℓ_1-penalized counterparts without magnitude prior. The difference stems from the fact that ℓ_2-regularization enforces a smooth susceptibility map whereas ℓ_1-penalty imposes a piece-wise constant image model that preserves sharp edges. The k-space views in Figs.8b and d also reveal that ℓ_1-regularization is more successful at estimating the undersampled content in the vicinity of the magic angle relative to ℓ_2-based QSM when no magnitude weighting is employed. To quantify the improvement, k-space energy of the reconstructed maps at the frequencies where the susceptibility kernel dampens the field map (frequencies $\boldsymbol{k}$ that satisfy $\mathbf{D}(k_x, k_y, k_z) \leq 0.25$) was computed. In this case, ℓ_1-based maps had 11 % higher energy than the ℓ_2-reconstruction. Considering only high frequency content at this threshold level ($\mathbf{D} \leq 0.25$ and $\sqrt{k_x^2+k_y^2+k_z^2} \geq 0.5 \cdot k_{max}$), ℓ_1-regularized map was found to have 69 % higher k-space energy. However, the difference between the two regularization approaches would be smaller when average susceptibility values are computed within anatomical boundaries of large iron-rich structures such as putamen and globus pallidus (16).

Magnitude-weighted regularization might mitigate this drawback of ℓ_2-based regularization through edge-aware smoothing. As seen in Figs.6a and b, the conspicuity of vessels has substantially increased when magnitude prior was involved in ℓ_2-regularization. When Figs. 6c and d are compared, benefit of magnitude prior can also be observed in ℓ_1-regularization, albeit to a smaller degree. Effect of edge-aware dipole inversion can also be observed in the susceptibility MIPs (Figs.7c and e) and in improved estimation of k-space content near the magic angle (Figs.8c and e).

Increased computational efficiency would have particular impact on processing large datasets. Recently introduced functional susceptibility mapping (6,24,25) involves reconstruction of a time series of three-dimensional phase images. The complete pipeline consisting of phase unwrapping, background removal, smoothing parameter value estimation and dipole inversion for 30 time frames at 1.5 mm isotropic resolution required 4 minutes when state-of-the-art phase processing and the proposed algorithm were combined. The reconstruction time would otherwise exceed one hour with the nonlinear Conjugate Gradient algorithm. As such, the proposed algorithm is expected to facilitate the investigation of the relation between functional blood oxygenation dependent (BOLD) contrast and changes in the underlying susceptibility distribution. To mitigate the low-spatial frequency background variation in fMRI phase images, a combination of phase images at different time-points can be utilized (e.g. by subtraction of the first time-point or of the average phase over time) and then χ can be computed (6). While TKD and ℓ_2-regularization are linear in χ, ℓ_1-regularization is a non-linear reconstruction technique. The resulting susceptibility maps will be independent of the order in which subtraction and dipole inversion are computed for TKD and ℓ_2-penalty, however the ordering is important for ℓ_1-constrained inversion because of the thresholding step.

Phase processing and dipole inversion for EPI data acquired at 7T constitutes a challenging problem, particularly due to imperfections associated with echo-planar k-space trajectory. While 1.5 mm isotropic voxel size (Figs.9 and 10) is high-resolution for fMRI, it is considerably lower than the resolution of the 3D GRE data acquired at 0.6 mm isotropic voxel size (Figs. 2–8). As such, the EPI volumes do not have the same level of spatial detail as the GRE images, however they are capable of representing variations in response to neural activity and physiological changes since they constitute a time-series dataset. Relative to the tissue phase images in Fig. 9, loss of spatial resolution can be observed in the ℓ_2- and ℓ_1-constrained time-series in Figs.10a and b. Although the smoothing parameters β and λ were selected with the L-curve method, the amount of regularization can also be tailored to a particular problem while using the L-curve parameter values as useful landmarks (e.g. a parameter 10% less than the L-curve selected value can be employed for reduced smoothing). Once the regularization parameter is selected on a single time point, the same value can be applied to the rest of the volumes in the time-series. This point constitutes the major computational difference between the high-resolution 3D GRE dataset and the functional QSM experiment.

Based on confounding effects of instrumental and physiological origin in the EPI phase signal (6), we further acknowledge that extracting functional information is a difficult task that requires stability over the time points. To quantify the stability in the raw unwrapped phase, tissue phase, and ℓ_2- and ℓ_1-regularized time-series, we report average time-SNR and maps of standard deviation over time in Figs.10c–f. The streaking artifacts visible in these maps stem from imperfect estimation of the relative phase offset of each coil. This can be mitigated by estimating the offset from a dual-echo acquisition (29). As a result of deconvolution, ℓ_2- and ℓ_1-based susceptibility series exhibit 33 % and 42 % increase in average standard deviation relative to the tissue phase. However, the time-SNR values in the susceptibility maps are larger by 42 % and 13 % when compared to times-SNR value of the

tissue phase, owing to the fact that the increase in the signal counterbalanced the increase in the noise. It can also be seen that the noise standard deviation for ℓ_1-based reconstruction has substantial spatial variation. Based on the nonlinear nature of regularization, smooth regions tend to remain below the ℓ_1-threshold which leads to small standard deviation over time. Further, a 5-fold reduction in time-SNR is observed between the raw unwrapped and the tissue phase. Since the tissue component is about an order of magnitude smaller than the background contribution, substantial reduction in the phase signal is expected when background phase is eliminated.

In addition to the regularization parameter λ that adjusts the contribution of the signal prior to the reconstructed susceptibility map, the proposed variable splitting formulation introduces a second parameter μ that weights the gradient consistency due to ($\mu \cdot \|\boldsymbol{y}-\mathbf{G}\boldsymbol{\chi}-\boldsymbol{\eta}\|_2^2$). While Ref. (21) shows that μ does not affect the solution but the speed of convergence, a suitable parameter value still needs to be selected. For the *in vivo* setting, we addressed the parameter identification problem by setting μ to the optimal ℓ_2-parameter β that was determined with the L-curve method. This heuristic selection was seen to yield favorable convergence speed. Based on the numerical phantom experiments detailed under Results section, the same reconstruction error was obtained when the value of μ varied within three orders of magnitude range. This points out that the same susceptibility map is obtained regardless of the value of μ. Regarding the convergence speed of the heuristically selected μ parameter, the experiments performed on the in vivo dataset demonstrated that using 10-times larger or 10-times smaller parameters lead to slower convergence. As such, the heuristically selected parameter is seen to have favorable convergence characteristics in practice.

Limitations

To compensate for the noise variation in the field map, a diagonal weighting $\mathbf{M}$ proportional to the image magnitude can be included in the data consistency term (10,12,13). With this refinement, the ℓ_2-constrained problem becomes,

$$min_{\boldsymbol{\chi}}\|\mathbf{M}(\mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\chi}-\boldsymbol{\phi})\|_2^2+\beta\cdot\|\mathbf{W}\mathbf{G}\boldsymbol{\chi}\|_2^2 \quad (21)$$

The optimizer of this expression is given by the solution of

$$(\mathbf{D}^T\mathbf{F}\mathbf{M}^2\mathbf{F}^{-1}\mathbf{D}+\beta\cdot\mathbf{E}^H\mathbf{F}\mathbf{W}^2\mathbf{F}^{-1}\mathbf{E})\mathbf{F}\boldsymbol{\chi}=\mathbf{D}^T\mathbf{F}\mathbf{M}^2\mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\phi} \quad (22)$$

This system needs to be solved iteratively, and the preconditioner ($\mathbf{D}^2 + \beta \cdot \mathbf{E}^2$) is expected to be less effective compared to its use without noise weighting $\mathbf{M}$. A similar system arises in the update formula for ℓ_1-based reconstruction,

$$(\mathbf{D}^T\mathbf{F}\mathbf{M}^2\mathbf{F}^{-1}\mathbf{D}+\mu\mathbf{E}^T\mathbf{F}\mathbf{W}^2\mathbf{F}^{-1}\mathbf{E})\mathbf{F}\boldsymbol{\chi}_{t+1}=\mathbf{D}^T\mathbf{F}\mathbf{M}^2\mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\phi}+\mu\mathbf{E}^T\mathbf{F}\mathbf{W}^T(\boldsymbol{y}_t-\boldsymbol{\eta}_t) \quad (23)$$

Employing efficient matrix factorization algorithms could potentially facilitate these more challenging matrix inversion problems (37,38).

Extensions

L-curve parameter estimation entails reconstruction with varying levels of regularization. Because each reconstruction is independent of the others, they could be performed in parallel for increased time efficiency.

The phase processing pipeline employed in the current work could be further refined. A regularized version of the SHARP filter was recently proposed to enhance the quality of the background phase removal (32,39). This improvement would however come at the cost of additional processing time, as this regularized formulation is solved iteratively. A second refinement would be to use a spatially varying SHARP kernel size, which would yield higher quality tissue phase inside the brain, and reduce the amount erosion that needs to be applied on the mask boundary (40). This improvement is included in the software package that accompanies this manuscript.

Conclusion

This work introduces a variable-splitting algorithm that reduces the processing time of ℓ_1-regularized QSM by 20 times relative to the conventional nonlinear Conjugate Gradient solver. Such efficient optimization also renders regularization parameter estimation with the L-curve method practical. Combined with state-of-the-art phase unwrapping and background removal techniques, the proposed algorithm comprises a pipeline that might facilitate clinical use of susceptibility mapping. This method is also extended to admit prior information derived from the magnitude signal for edge-aware regularization. The developed fast dipole inversion methods are expected to facilitate the investigation of the relation between the BOLD signal and the underlying tissue susceptibility changes by reconstructing four-dimensional time-series datasets in feasible time.

References

1. Langkammer C, Schweser F, Krebs N, et al. Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. Neuro Image. 2012; 62:1593–9.10.1016/j.neuroimage.2012.05.049 [PubMed: 22634862]
2. Zheng W, Nichol H, Liu S, Cheng Y, Haacke E. Measuring iron in the brain using quantitative susceptibility mapping and X-ray fluorescence imaging. Neuro Image. 201310.1016/j.neuroimage.2013.04.022
3. Fan, AP.; Bilgic, B.; Benner, T.; Rosen, BR.; Adalsteinsson, E. Regularized Quantitative Susceptibility Mapping for Phase-based Regional Oxygen Metabolism (PROM) at 7T. Proc Intl Soc Mag Reson Med; 2011; Montreal, Canada. p. 7942
4. Fan AP, Benner T, Bolar DS, Rosen BR, Adalsteinsson E. Phase-based regional oxygen metabolism (PROM) using MRI. Magnetic Resonance in Medicine. 2012; 67:669–78.10.1002/mrm.23050 [PubMed: 21713981]
5. Haacke E, Tang J, Neelavalli J, Cheng Y. Susceptibility mapping as a means to visualize veins and quantify oxygen saturation. Journal of Magnetic Resonance Imaging. 2010; 32:663–676.10.1002/jmri.22276 [PubMed: 20815065]

6. Bianciardi M, van Gelderen P, Duyn JH. Investigation of BOLD fMRI resonance frequency shifts and quantitative susceptibility changes at 7 T. Human brain mapping. 2013 In Press. 10.1002/hbm.22320
7. Langkammer C, Liu T, Khalil M, Enzinger C, Jehna M, Fuchs S, Fazekas F, Wang Y, Ropele S. Quantitative Susceptibility Mapping in Multiple Sclerosis. Radiology. 2013; 267:551–9.10.1148/radiol.12120707 [PubMed: 23315661]
8. Schweser F, Deistung A, Lehr B, Reichenbach J. Differentiation between diamagnetic and paramagnetic cerebral lesions based on magnetic susceptibility mapping. Medical physics. 2010; 37:5165–78. [PubMed: 21089750]
9. Marques JP, Bowtell R. Application of a Fourier-based method for rapid calculation of field inhomogeneity due to spatial variation of magnetic susceptibility. Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering. 2005; 25B:65–78.10.1002/cmr.b.20034
10. De Rochefort L, Liu T, Kressler B, Liu J, Spincemaille P, Lebon V, Wu J, Wang Y. Quantitative susceptibility map reconstruction from MR phase data using bayesian regularization: validation and application to brain imaging. Magnetic Resonance in Medicine. 2010; 63:194–206.10.1002/mrm.22187 [PubMed: 19953507]
11. Liu T, Liu J, de Rochefort L, Spincemaille P, Khalidov I, Ledoux JR, Wang Y. Morphology enabled dipole inversion (MEDI) from a single-angle acquisition: comparison with COSMOS in human brain imaging. Magnetic Resonance in Medicine. 2011; 66:777–83.10.1002/mrm.22816 [PubMed: 21465541]
12. 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. Neuro Image. 2012; 59:2560–8.10.1016/j.neuroimage.2011.08.082 [PubMed: 21925276]
13. 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. Magnetic Resonance in Medicine. 2009; 61:196–204.10.1002/mrm.21828 [PubMed: 19097205]
14. Schweser F, Deistung A, Lehr BW, Reichenbach JR. Quantitative imaging of intrinsic magnetic tissue properties using MRI signal phase: an approach to in vivo brain iron metabolism? Neuro Image. 2011; 54:2789–807.10.1016/j.neuroimage.2010.10.070 [PubMed: 21040794]
15. Wharton S, Bowtell R. Whole-brain susceptibility mapping at high field: a comparison of multiple- and single-orientation methods. Neuro Image. 2010; 53:515–25.10.1016/j.neuroimage.2010.06.070 [PubMed: 20615474]
16. Bilgic B, Pfefferbaum A, Rohlfing T, Sullivan EV, Adalsteinsson E. MRI estimates of brain iron concentration in normal aging using quantitative susceptibility mapping. Neuro Image. 2012; 59:2625–35.10.1016/j.neuroimage.2011.08.077 [PubMed: 21925274]
17. Wu B, Li W, Guidon A, Liu C. Whole brain susceptibility mapping using compressed sensing. Magnetic Resonance in Medicine. 2012; 67:137–47.10.1002/mrm.23000 [PubMed: 21671269]
18. Hansen, PC. The L-Curve and its Use in the Numerical Treatment of Inverse Problems. WIT Press; Southampton: 2000. p. 119-142.
19. Morozov V. On the solution of functional equations by the method of regularization. Soviet Math Dokl. 1966; 7:414–417.
20. Bilgic, B.; Chatnuntawech, I.; Fan, AP.; Adalsteinsson, E. Regularized QSM in Seconds. Proc. Intl. Soc. Mag. Reson. Med; 2013; Salt Lake City, USA. p. 168
21. Goldstein T, Osher S. The split Bregman method for L1-regularized problems. SIAM Journal on Imaging Sciences. 2009; 2:323–343.
22. Li W, Wu B, Avram AV, Liu C. Magnetic susceptibility anisotropy of human brain in vivo and its molecular underpinnings. Neuro Image. 2012; 59:2088–97.10.1016/j.neuroimage.2011.10.038 [PubMed: 22036681]
23. Chen Z, Calhoun V. Computed Inverse Resonance Imaging for Magnetic Susceptibility Map Reconstruction. Journal of Computer Assisted Tomography. 2012; 36:265–274. [PubMed: 22446372]

24. Balla, DZ.; Sanchez-Panchuelo, RM.; Wharton, S.; Hagberg, GE.; Scheffler, K.; Francis, ST.; Bowtell, RW. Functional Quantitative Susceptibility Mapping (fQSM). Proc. Intl. Soc. Mag. Reson. Med; 2012; Melbourne, Australia. p. 325
25. Balla, DZ.; Sanchez-Panchuelo, RM.; Wharton, S.; Hagberg, GE.; Scheffler, K.; Francis, ST.; Bowtell, RW. Experimental investigation of the relation between gradient echo BOLD fMRI contrast and underlying susceptibility changes at 7T. Proc. Intl. Soc. Mag. Reson. Med; 2013; Salt Lake City, USA. p. 300
26. Liu T, Xu W, Spincemaille P, Avestimehr AS, Wang Y. Accuracy of the morphology enabled dipole inversion (MEDI) algorithm for quantitative susceptibility mapping in MRI. Medical Imaging, IEEE Transactions on. 2012; 31:816–824.
27. Duyn J, van Gelderen P, Tie-Qiang L, de Zwart JA, Koretsky AP, Fukunaga M. High-field MRI of brain cortical substructure based on signal phase. Proceedings of the National Academy of Sciences. 2007; 104:11796–11801.
28. Deistung A, Dittrich E, Sedlacik J, Rauscher A, Reichenbach JR. ToF-SWI: Simultaneous time of flight and fully flow compensated susceptibility weighted imaging. Journal of Magnetic Resonance Imaging. 2009; 29:1478–1484.10.1002/jmri.21673 [PubMed: 19472425]
29. Robinson S, Grabner G, Witoszynskyj S, Trattnig S. Combining phase images from multi-channel RF coils using 3D phase offset maps derived from a dual-echo scan. Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine. 2011; 65:1638–48.10.1002/mrm.22753
30. Hammond KE, Lupo JM, Xu D, Metcalf M, Kelley DAC, Pelletier D, Chang SM, Mukherjee P, Vigneron DB, Nelson SJ. Development of a robust method for generating 7. 0 T multichannel phase images of the brain with application to normal volunteers and patients with neurological diseases. Neuro Image. 2008; 39:1682–92.10.1016/j.neuroimage.2007.10.037 [PubMed: 18096412]
31. Smith S. Fast robust automated brain extraction. Human brain mapping. 2002; 17:143–155.10.1002/hbm.10062 [PubMed: 12391568]
32. Sun, H.; Wilman, AH. Susceptibility Mapping Using Regularization Enabled Harmonic Artifact Removal. Proc. Intl. Soc. Mag. Reson. Med; 2013; Salt Lake City, USA. p. 169
33. Keil, B.; Triantafyllou, C.; Hamm, M.; Wald, LL. Design optimization of a 32-channel head coil at 7 T. Proc. Intl. Soc. Mag. Reson. Med; 2010; Stockholm, Sweden. p. 1493
34. Montesinos, P.; Abascal, JJ.; Chamorro, J.; Chavarrias, C.; Benito, M.; Vaquero, JJ.; Desco, M. High-resolution dynamic cardiac MRI on small animals using reconstruction based on Split Bregman methodology. Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC); 2011. p. 3462-3464.
35. Schweser F, Deistung A, Sommer K, Reichenbach JR. Toward online reconstruction of quantitative susceptibility maps: Superfast dipole inversion. Magnetic Resonance in Medicine. 2012; 69:1581–93.10.1002/mrm.24405
36. Shmueli K, De Zwart J. Magnetic susceptibility mapping of brain tissue in vivo using MRI phase data. Magnetic Resonance in Medicine. 2009; 62:1510–1.10.1002/mrm.22135 [PubMed: 19859937]
37. Xia J, Chandrasekaran S, Gu M, Xiaoye S. Fast algorithms for hierarchically semiseparable matrices. Numerical Linear Algebra with Applications. 2010; 17:953–976.10.1002/nla.691
38. Xia J, Chandrasekaran S, Gu M, Li X. Superfast multifrontal method for large structured linear systems of equations. SIAM Journal on Matrix Analysis and Applications. 2009; 31:1382–1411.
39. Sun H, Wilman A. Background field removal using spherical mean value filtering and Tikhonov regularization. Magnetic Resonance in Medicine. 201310.1002/mrm.24765
40. Wu B, Li W, Avram AV, Gho S-M, Liu C. Fast and tissue-optimized mapping of magnetic susceptibility and T2* with multi-echo and multi-shot spirals. Neuro Image. 2012; 59:297–305.10.1016/j.neuroimage.2011.07.019 [PubMed: 21784162]

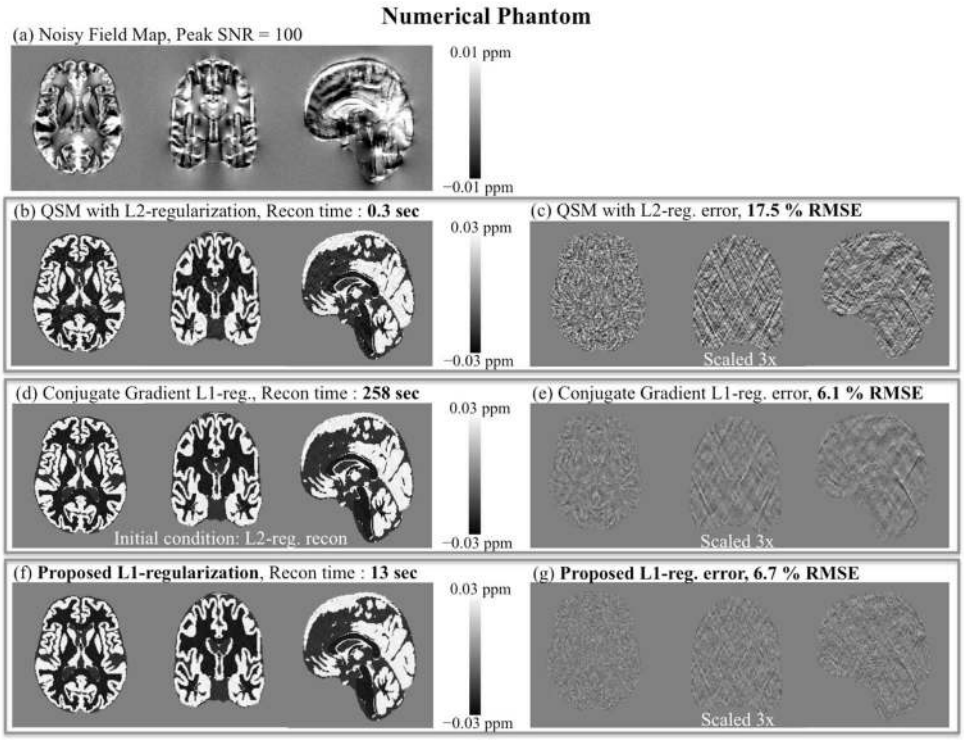

Fig. 1.

QSM with numerical phantom. RMSE values are computed relative to the known, ground-truth susceptibility map. A field map was simulated using the true susceptibility, to which noise with PSNR=100 was added to obtain the noisy field map in (a). QSM with closed-form L2-regularization shown in (b) took 0.3 seconds, and the reconstruction error (c) was 17.5 %. L1-regularized Conjugate Gradient reconstruction took 258 seconds (d), and the error (e) was 6.1 %. Proposed L1-constrained QSM was completed in 13 seconds (f), with an RMSE of 6.7 % (g).

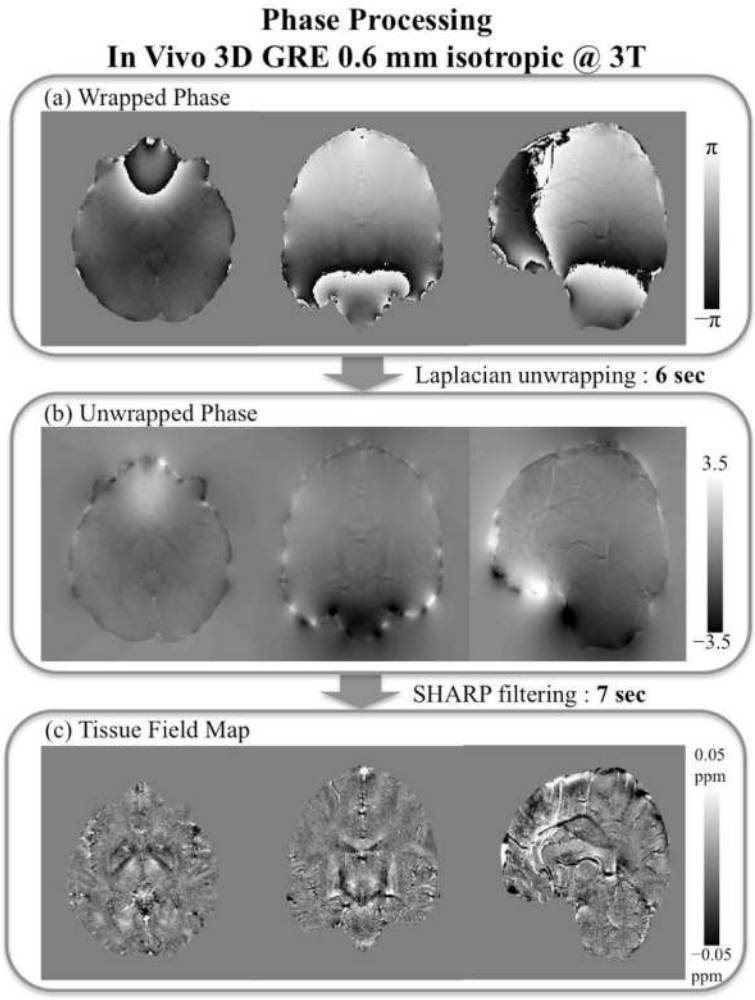

Fig. 2.

Phase processing steps for in vivo 3D GRE data at 0.6 mm resolution. Starting from the coil-combined, wrapped phase in (a), unwrapped phase data are obtained with Laplacian unwrapping (b) in 6 seconds. Further processing with SHARP filtering yielded the tissue field map (c) in 7 seconds.

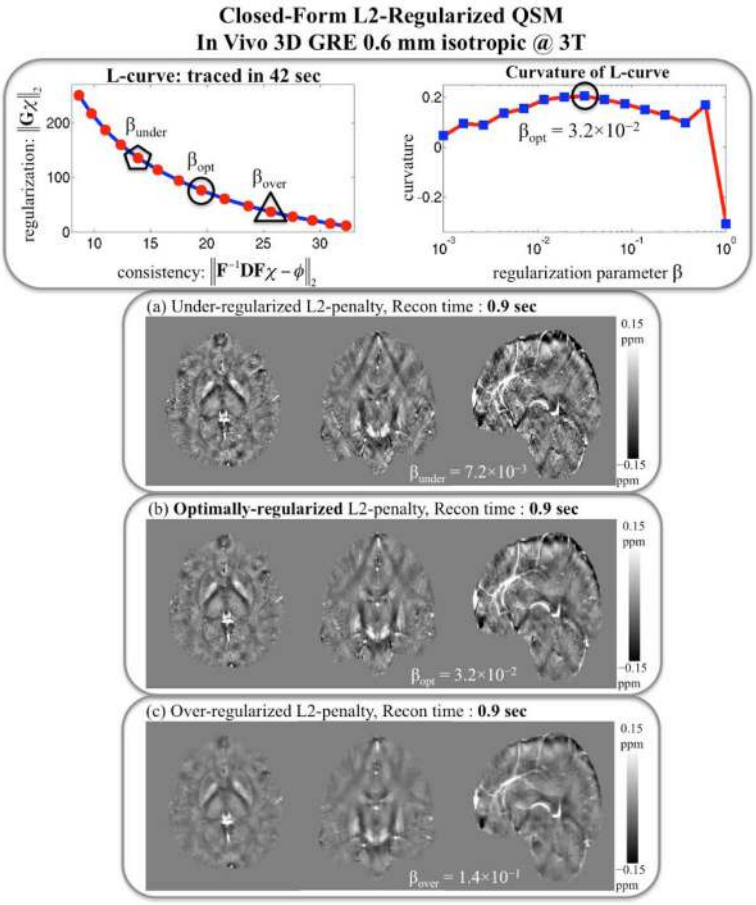

Fig. 3.

Closed-form L2-constrained reconstruction for 3D GRE. Upper panel: L-curve is traced in 42 seconds, and the parameter value that maximized the curvature was $\beta = 3.2 \cdot 10^{-2}$, corresponding to the optimal level of regularization. In (a), (b) and (c), under-, optimally- and over-regularized susceptibility maps are depicted. Each reconstruction took 0.9 seconds of computation time.

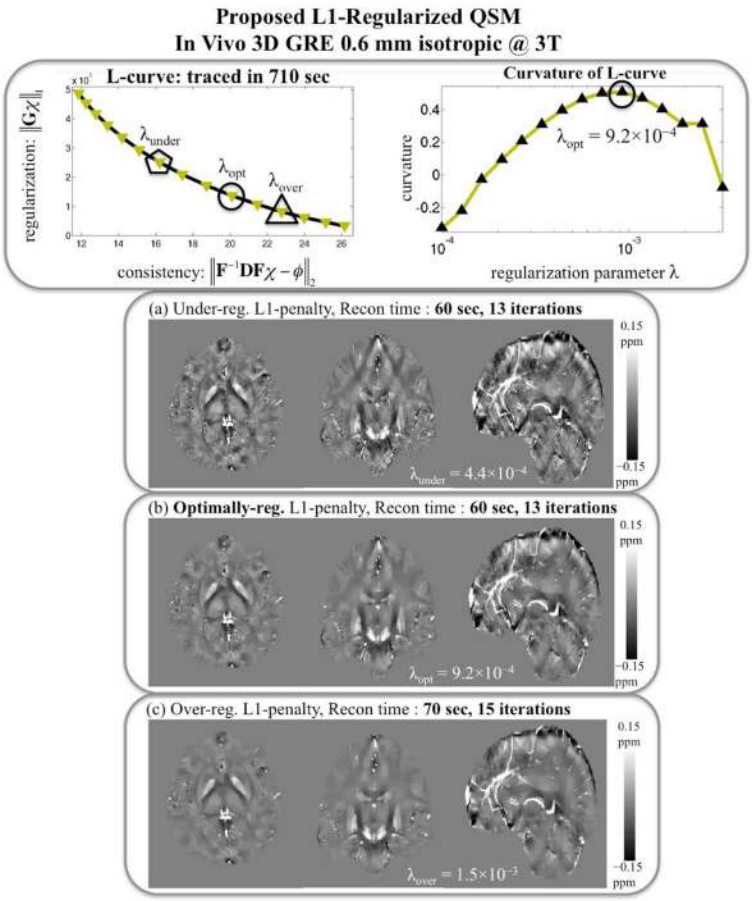

Fig. 4.

Proposed L1-constrained reconstruction for 3D GRE. Upper panel: L-curve is traced in 710 seconds, and the parameter value that maximized the curvature was $\lambda = 9.2 \cdot 10^{-4}$, corresponding to the optimal level of regularization. In (a), (b) and (c), under-, optimally- and over-regularized susceptibility maps are depicted. Under- and optimally-regularized reconstructions took 60 seconds and 13 iterations to converge, while optimization took was 70 seconds and 15 iterations for the over-regularized case.

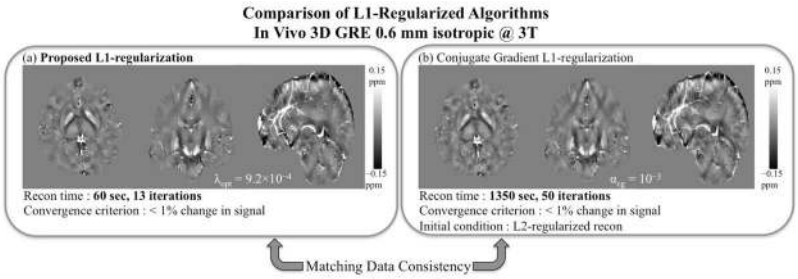

Fig. 5.

Comparison of L1-regularized dipole inversion methods for in vivo 3D GRE. Proposed algorithm in (a) converged in 60 seconds and 13 iterations, while it took 1350 seconds and 50 iterations for the Conjugate Gradient algorithm to finish (b).

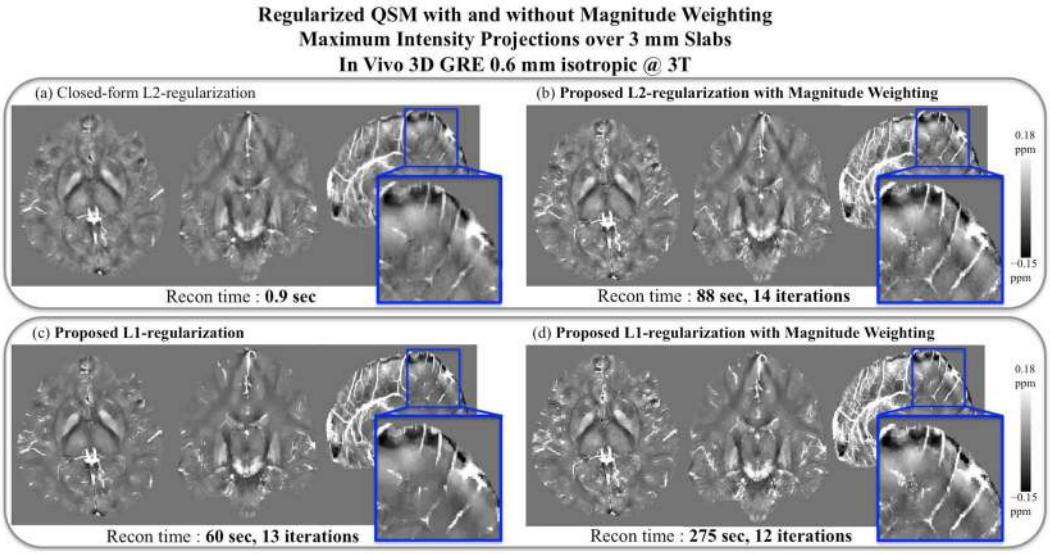

Fig. 6.

L2- and L1-regularized QSM with and without magnitude prior. Compared to closed-form reconstruction in (a) that is completed in 0.9 seconds, magnitude weighted L2-regularization in (b) requires 88 seconds of processing while increasing conspicuity of high-frequency structures like vessels. Relative to the proposed L1-based method in (c), inclusion of magnitude prior in (d) is computationally more demanding, requiring 275 seconds of reconstruction.

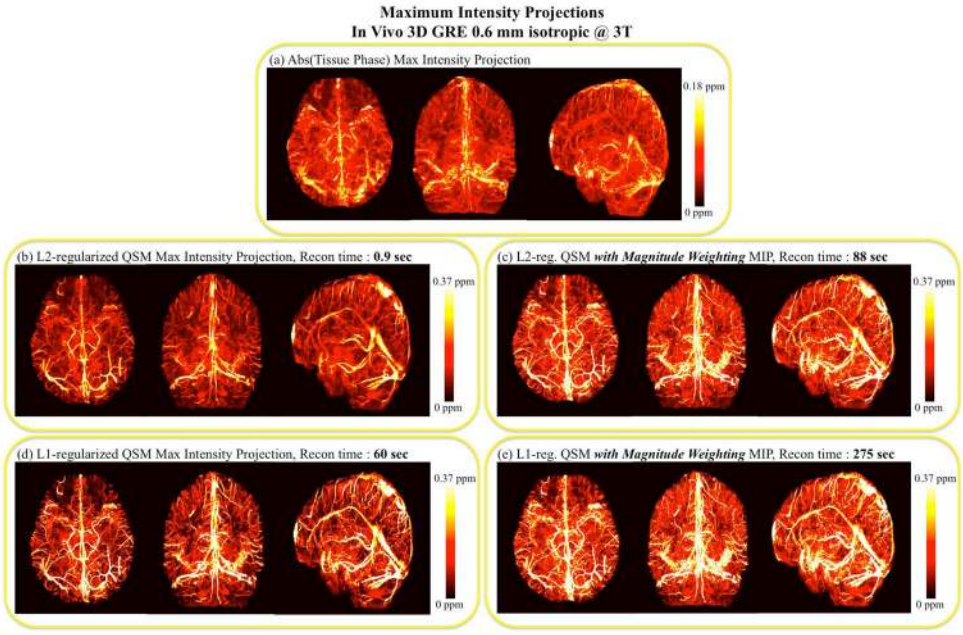

Fig. 7.

Maximum intensity projections (MIPs) of in vivo 3D GRE dataset. Tissue phase MIP is shown in (a), and closed-form L2-based susceptibility map is depicted in (b). Projection for the proposed L2-regularized QSM with magnitude prior is given in (c), and L1-based reconstruction without (d) and with magnitude weighting is shown in (e). Note the increase in the vessel susceptibility values estimated with the methods that utilize magnitude prior.

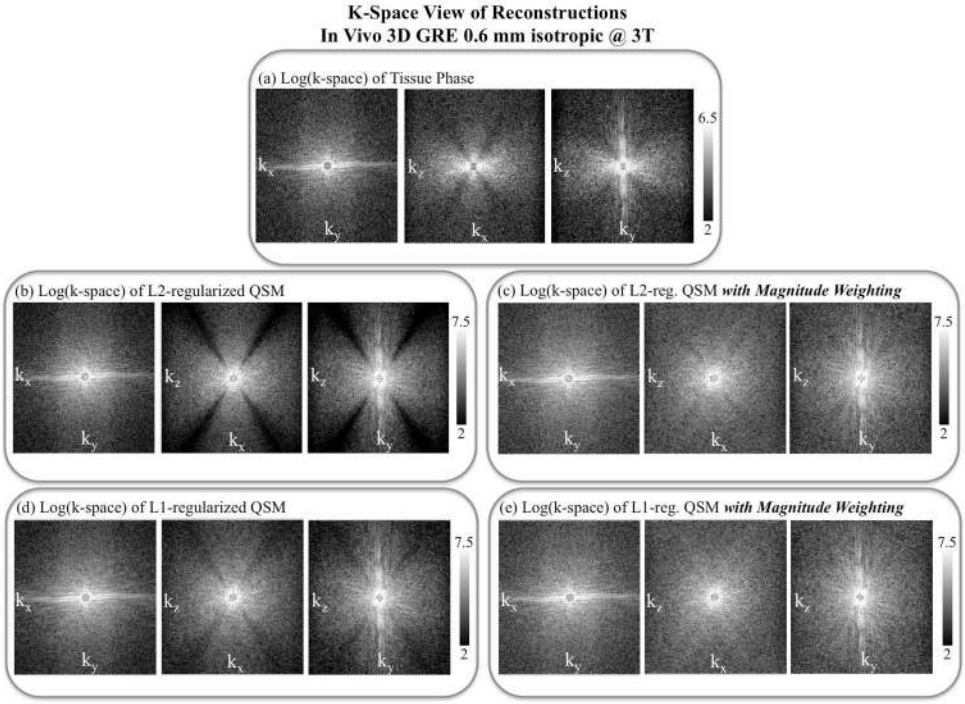

Fig. 8.

K-space views for tissue phase (a), closed-form QSM (b), proposed L2-regularization with magnitude weighting (c), L1-based reconstruction without (d), and with magnitude prior (e). Note the increase in the k-space content near the magic angle for the methods that utilize magnitude prior.

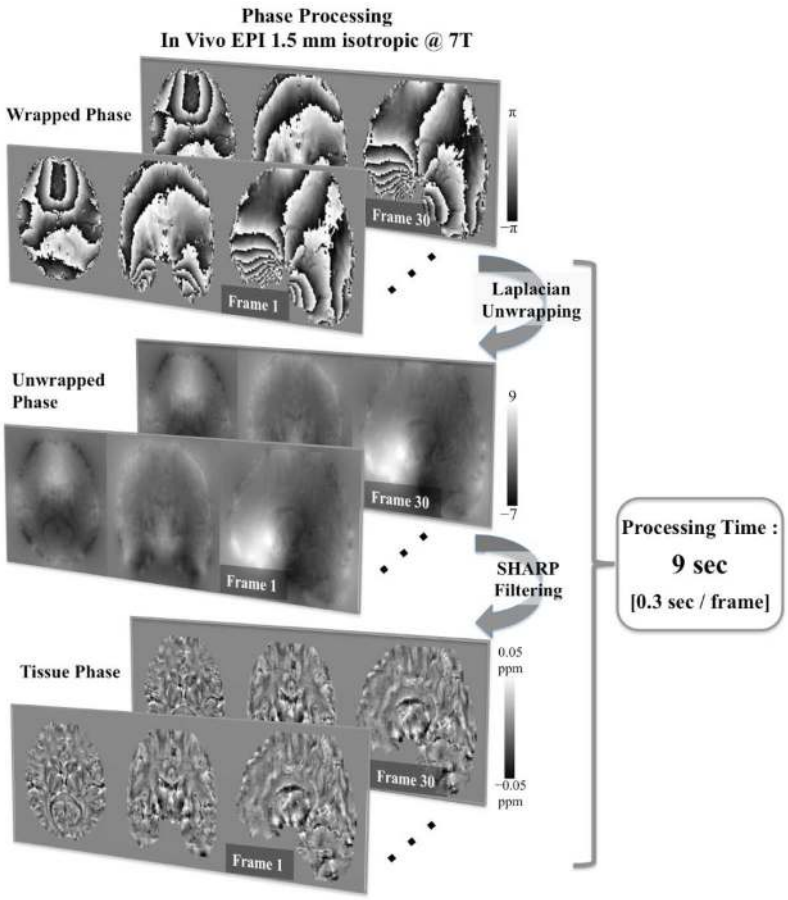

Fig. 9.

Phase processing steps for in vivo EPI at 1.5 mm isotropic resolution acquired as a time-series (frames 1 to 30). Starting from the coil-combined wrapped phase, application of Laplacian unwrapping and SHARP filtering took 9 seconds for the 30 frames, corresponding to 0.3 seconds/frame processing time.

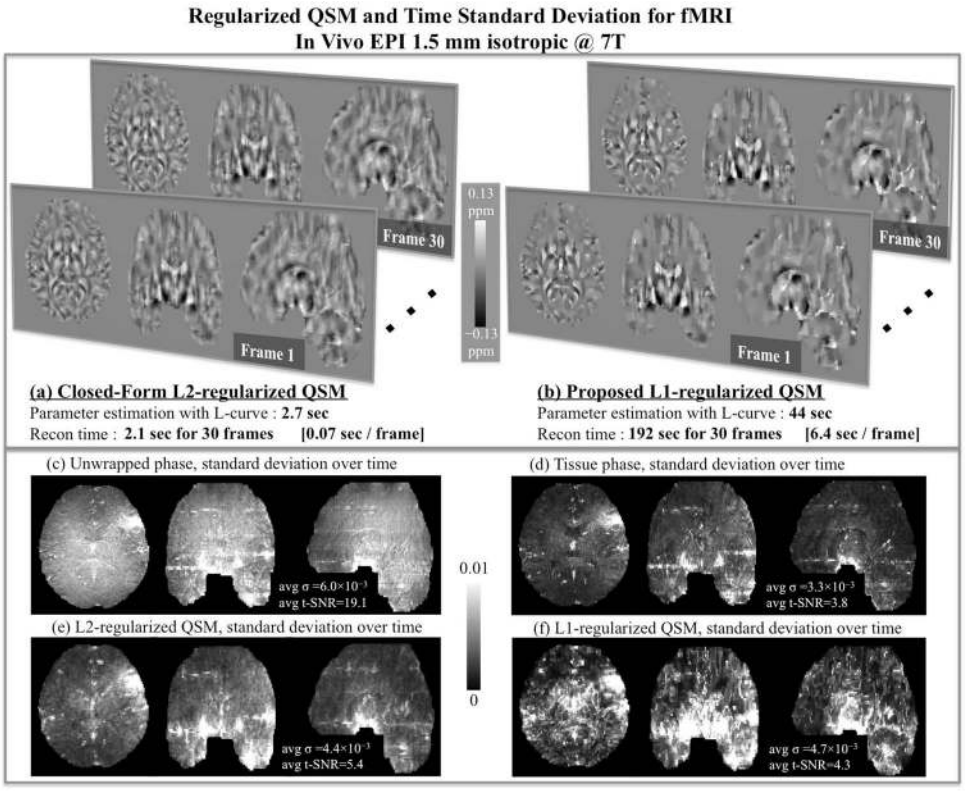

Fig. 10.

Reconstruction of 30 frames of EPI data with the closed-form L2-regularized QSM shown in (a) was completed in 2.1 seconds, corresponding to 0.07 seconds/frame speed. Using the proposed L1-based method shown in (b), the reconstruction time was 192 seconds for the 30 frames, yielding a processing speed of 6.4 seconds/frame. L-curve parameter estimation took 2.7 seconds for L2- and 44 seconds for L1-constrained reconstruction. Standard deviation maps of phase and susceptibility time-series are depicted in (c)–(f). Raw unwrapped phase in (c) has a standard deviation of $6 \cdot 10^{-3}$ over time, and a time-SNR of 19.1 averaged inside the brain mask. For tissue phase shown in (d), these values were $\sigma = 3.3 \cdot 10^{-3}$ and t-SNR=3.8. L2-regularized susceptibility time-series had $\sigma = 4.4 \cdot 10^{-3}$ and t-SNR=5.4 (e), while L1-based reconstruction returned $\sigma = 4.7 \cdot 10^{-3}$ and t-SNR=4.3 (f).

Table 1

QSM reconstruction algorithms and related regularization parameters

QSM Algorithm	Parameter	Numerical Phantom	In Vivo 3D GRE	In Vivo EPI
Closed-form ℓ_2-regularization	β (ℓ_2 penalty)	$2.2 \cdot 10^{-4}$	$3.2 \cdot 10^{-2}$	$3.2 \cdot 10^{-2}$
Nonlinear CG ℓ_1-regularization	α (ℓ_1 penalty)	$1.5 \cdot 10^{-5}$	$1.0 \cdot 10^{-3}$	–
Proposed ℓ_1-regularization	λ (ℓ_1 penalty)	$1.0 \cdot 10^{-5}$	$9.2 \cdot 10^{-4}$	$9.2 \cdot 10^{-4}$
	μ (gradient consistency)	$2.2 \cdot 10^{-4}$	$3.2 \cdot 10^{-2}$	$3.2 \cdot 10^{-2}$