Nonlinear dipole inversion (NDI) enables robust quantitative susceptibility mapping (QSM)

Daniel Polak\textsuperscript{1,2,3} | Itthi Chatnuntawech\textsuperscript{4} | Jaeyeon Yoon\textsuperscript{5} | Siddharth Srinivasan Iyer\textsuperscript{2,6} | Carlos Milovic\textsuperscript{7} | Jongho Lee\textsuperscript{5} | Peter Bachert\textsuperscript{1,8} | Elfar Adalsteinsson\textsuperscript{6} | Kawin Setsompop\textsuperscript{2,9,10} | Berkin Bilgic\textsuperscript{2,9,10}

1Department of Physics and Astronomy, Heidelberg University, Heidelberg, Germany
2Department of Radiology, A. A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, Massachusetts, USA
3Siemens Healthcare GmbH, Erlangen, Germany
4National Science and Technology Development Agency, National Nanotechnology Center, Pathum Thani, Thailand
5Laboratory for Imaging Science and Technology, Department of Electrical and Computer Engineering, Seoul National University, Seoul, South Korea
6Department of Electronic Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
7Department of Electrical Engineering, Pontificia Universidad Catolica de Chile, Santiago, Chile
8Medical Physics in Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
9Harvard Medical School, Boston, Massachusetts, USA
10Harvard-MIT Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Correspondence
Itthi Chatnuntawech, National Nanotechnology Center, National Science and Technology Development Agency, Pathum Thani, Thailand.
Email: itthi.cha@nanotec.or.th

Funding information
Foundation for the National Institutes of Health, Grant/Award Numbers:
P41EB015896, R01EB019437, R01EB020613, S10RR019254, S10RR019307, S10RR023043, S10RR023401, U01EB025162

High-quality Quantitative Susceptibility Mapping (QSM) with Nonlinear Dipole Inversion (NDI) is developed with pre-determined regularization while matching the image quality of state-of-the-art reconstruction techniques and avoiding over-smoothing that these techniques often suffer from. NDI is flexible enough to allow for reconstruction from an arbitrary number of head orientations and outperforms COSMOS even when using as few as 1-direction data. This is made possible by a nonlinear forward-model that uses the magnitude as an effective prior, for which we derived a simple gradient descent update rule. We synergistically combine this physics-model with a Variational Network (VN) to leverage the power of deep learning in the VaNDI algorithm. This technique adopts the simple gradient descent rule from NDI and learns the network parameters during training, hence requires no additional parameter tuning. Further, we evaluate NDI at 7 T using highly accelerated Wave-CAIPI acquisitions at 0.5 mm isotropic resolution and demonstrate high-quality QSM from as few as 2-direction data.

KEYWORDS
deep learning, nonlinear inversion, quantitative susceptibility mapping

Abbreviations: CNN, convolutional neural network; MEDI, morphology enabled dipole inversion; NDI, nonlinear dipole inversion; QSM, quantitative susceptibility mapping; SMV, spherical mean value; TGV, total generalized variation; TKD, truncated k-space division; TV, total variation; VaNDI, variational nonlinear dipole inversion; VN, variational network; w.r.t, with respect to.
Quantitative Susceptibility Mapping (QSM) provides exquisite gray/white matter contrast\(^1\) and enables accurate quantification of iron in the brain.\(^2\) It is also utilized to differentiate dia- and paramagnetic sources of contrast,\(^3\) detect tissue changes related to neuro-degenerative diseases\(^4-6\) and estimate vessel oxygenation.\(^7-9\) However, it entails a difficult image reconstruction pipeline with several pre-processing steps, which are briefly summarized at the beginning of this contribution. This work will focus on the dipole inversion step, which aims to estimate the desired magnetic susceptibility from the acquired gradient echo phase information. Several benchmark and state-of-the-art techniques are described, and finally a new method for robust dipole inversion is introduced which mitigates some of the drawbacks of previous approaches.

Typically, the first step of the QSM reconstruction is the coil combination\(^10-13\) of multi-channel receive data. Hereby, phase offsets between the different coil images are estimated to prevent destructive interference between complex signals. The resulting coil-combined phase images are initially wrapped into an interval of \([-\pi, \pi]\). To recover the underlying phase distribution of interest, a spatial image unwrapping technique (path-following\(^14-17\); Laplacian\(^18-20\)) is applied, which usually relies on the assumption that the phase signal varies slowly from voxel to voxel. After unwrapping, the phase images are dominated by background phase effects which are commonly one to two orders of magnitude larger than those caused by the desired tissue susceptibility.\(^21-23\) Several filtering techniques have been proposed\(^24-28\) to remove unwanted phase contributions such as those caused by inhomogeneities of the static main magnetic field, macroscopic currents (MRI shim coils), or magnetic susceptibility variations outside the ROI (e.g. air-tissue interfaces).\(^29,30\) In a final step, the tissue phase \(\phi\) needs to be de-convolved with a known dipole kernel \(d(k) = 1/3 - k^2/k^2\) to obtain the desired susceptibility \(\chi\). Since the dipole kernel is not invertible, this inverse problem \(\phi(k) = d(k)\chi(k)\) cannot be solved using a simple division in k-space but requires more sophisticated reconstruction techniques.

Truncated k-space Division (TKD)\(^31\) inverts the dipole kernel directly \(\chi(k) = \tilde{d}(k)\phi(k)\) where small absolute values in \(d(k)\) are replaced by a constant number.

\[
\tilde{d}(k) = \begin{cases} d(k)^{-1} & \text{if } |d(k)| > \delta \\ \text{sgn}(d(k)) \cdot \delta^{-1} & \text{otherwise} \end{cases}
\]

However, the modification of the dipole kernel may result in systematic underestimation of the tissue susceptibility\(^32\) as well as streaking artifacts and noise amplification.

In COSMOS,\(^33\) gradient echo data are acquired under multiple head orientations \((r \in \{2 \ldots N\})\) with respect to the main magnetic field. This allows the optimization problem to be formulated in matrix form.

\[
\begin{pmatrix} d_1(k) \\ \vdots \\ d_N(k) \end{pmatrix} \chi(k) = \begin{pmatrix} \phi_1(k) \\ \vdots \\ \phi_N(k) \end{pmatrix}
\]

It also admits the closed form solution \(\chi = \left(\sum_i d_i^\dagger d_i\right)^{-1} \sum_i d_i^\dagger \phi_i\). With increasing number of head orientations, the conditioning of \(\left(\sum_i d_i^\dagger d_i\right)^{-1}\) improves since the dipole kernel also rotates and diminishing values around the magic angle do not overlap. This enables high image quality but comes at the cost of long scan time as multi-orientation data need to be acquired. This drawback was partly mitigated using fast imaging techniques such as EPI\(^24-26\) or Wave-CAIPI\(^37\); nevertheless, unnatural head positions/orientations remain a challenge for clinical translation.

Over the last decade, several single-orientation QSM reconstructions were proposed where additional regularization \(\|R(\chi)\|_\ell\) is used to improve the image quality. Commonly encountered regularizers utilize \(\ell_1\) or \(\ell_2\) penalties (such as in MEDI\(^4\) or L2\(^38\)), which are either applied on the image itself or on its representation using a transform. Note that the optimization problem shown below is posed in image space where \(D = F^\dagger dF\) applies the forward and inverse Fourier transform \(F\).

\[
\min_{\hat{\chi}} \|D\hat{\chi} - \hat{\phi}\|_2^2 + \lambda \|R(\hat{\chi})\|_\ell
\]

However, this formulation assumes that the linear susceptibility-to-field relationship is governed by Gaussian noise, whereas the phase noise distribution deviates from this especially in low-SNR regions.\(^39\) This was recognized in nonlinear-MEDI (NMEDI),\(^40\) where a nonlinear fidelity term was utilized.

\[
\min_{\hat{\chi}} \|W(e^{D\hat{\chi}} - e^{\hat{\phi}}\|_2^2 + \lambda \|MG\hat{\chi}\|_1
\]
Here the magnitude $W$ serves as a noise-weighting factor as well as allowing the derivation of a binary mask $M$ that weights the gradient $G$. This approach efficiently mitigates artifacts and improves the image appearance. However, the image quality strongly depends on the choice of the regularization parameter $\lambda$ which balances accuracy (data consistency) vs. image smoothness. Another drawback is the time-consuming reconstruction using complicated optimizers. This issue was addressed by the recently proposed FANSI algorithm\textsuperscript{41} which minimizes the same objective function as in NMEDI but utilizes a different optimization technique (Split Bregman method\textsuperscript{42} and ADMM\textsuperscript{43}). This allowed the regularization penalty and non-linear data fidelity to be converted into simpler decoupled problems (variable splitting), which can be solved more rapidly providing up to 10-fold computational speed-up. However, this acceleration comes at the cost of additional regularization parameters that need to be tuned manually (a total of four parameters are needed).

Alternative approaches employ Total Variation (TV) to promote image smoothness and reduce streaking artifacts.\textsuperscript{44} However, as TV only takes the first derivative into account, it neglects higher order smoothness and hence assumes that images are piecewise constant, which may lead to over-smoothing and unnatural image appearance. Total Generalized Variation (TGV) lifts this assumption by balancing both first and second derivatives, which is demonstrated to improve the image quality and prevent staircase artifacts.\textsuperscript{34}

Further improvement in image quality was achieved using single-step reconstruction algorithms\textsuperscript{45} which were proposed to mitigate potential error propagation between subsequent procedures along the QSM pipeline. As demonstrated in,\textsuperscript{45} operators for Laplacian unwrapping and spherical mean value (SMV) background filtering can be directly integrated into the optimization problem. While this increases the computational footprint, it may further reduce reconstruction errors when compared to multi-step reconstruction algorithms.

Recent advances in deep learning have gained widespread attention in the MRI research community. Convolutional neural networks (CNN) were trained to perform the deconvolution based on single-orientation phase data (QSMNet\textsuperscript{46}) and provided similar quality as multi-orientation COSMOS reconstructions. DeepQSM\textsuperscript{47} further demonstrated that the mathematical principle of dipole inversion can be learned entirely using synthetic images, and this network generalized to unseen in vivo data. As such, DeepQSM could potentially circumvent the demand for large amounts of patient training data and find important applications where ground-truth COSMOS data are not available for training such as in abdominal imaging.

In this contribution, we develop a simple gradient descent optimizer - Nonlinear Dipole Inversion (NDI) - and demonstrate how magnitude weighting and nonlinear formulation act as inherent priors, enabling high quality QSM with pre-determined regularization (no parameter tuning). We then expand NDI to learn variational regularizers from training data to further improve the image quality. Ultimately, we leverage Wave-CAIPI encoding to acquire highly accelerated high-resolution data at 7 T and evaluate the performance of NDI at 0.5 mm isotropic resolution.

\section*{Code/data: https://bit.ly/2RHeiF0}

\section*{2 \hspace{1em} METHOD}

\subsection*{2.1 \hspace{1em} Nonlinear dipole inversion (NDI)}

NDI is based on the nonlinear-MEDI (NMEDI)\textsuperscript{40} approach, but additional regularization terms are entirely removed and magnitude weighting and nonlinear formulation are exploited as inherent regularizers for the NDI reconstruction via the following cost function:

\[ f(\hat{x}) = \| W (e^{\tilde{D}_r \hat{x}} - e^{\tilde{\phi}_r}) \|_2^2 \]

This allows an analytical derivation of the gradient $\nabla_{\hat{x}} f(\hat{x})$ (see\textsuperscript{40} and Appendix for details) and the application of gradient descent optimization.

\[ \nabla_{\hat{x}} f(\hat{x}) = 2 D^T W^T W \sin (D \hat{x} - \tilde{\phi}_r) \]

With this, the $t^{th}$ update (iteration) of the reconstruction becomes

\[ \hat{x}^{t+1} = \hat{x}^t - 2 \sum_{r=1}^{N} D^T W^T W_r \sin (D \hat{x}^t - \tilde{\phi}_r) \]

where we generalized the formula for multi-orientation reconstruction from $N$-directions, with $W_r$, $\tilde{\phi}_r$, and $D_r$ denoting the magnitude, tissue phase and dipole kernel belonging to the $r^{th}$ head orientation. Also, this framework can be easily expanded to allow for Tikhonov regularization by subtracting $2\lambda \hat{x}^t$ from the above equation.
2.2 | Data acquisition and preparation

We used the QSMNet dataset\(^46\) where 3D-GRE data were acquired at 3 T from nine subjects using five head orientations with 1 mm isotropic resolution, 256x224x192 matrix, TE/TR = 25/33 ms, flip-angle = 15\(^\circ\), bandwidth = 100 Hz/px, R = 2x2 GRAPPA acceleration and TA = 5:46 min per orientation.

On a 7 T research system (Siemens Healthcare, Erlangen, Germany), we acquired additional 3D-GRE data at 0.5 mm isotropic resolution from one healthy volunteer using a prototype Wave-CAIPI\(^37\) sequence (3 head orientations, 480x480x360 matrix, TE/TR = 19/29 ms, flip angle = 25\(^\circ\), bandwidth = 100 Hz/px, R = 5x3 acceleration, acquisition time TA = 5:13 min per orientation). A custom tight-fitting 31-channel head coil\(^48\) (non-product) was used to achieve high-quality imaging; however, this limited the feasible head orientations to shallow angles (0\(^\circ\), 7\(^\circ\), 13\(^\circ\)). Additional low-resolution GRE reference scans were acquired for each head orientation to compute coil sensitivity maps using ESPIRiT.\(^49\) The parallel imaging reconstruction was performed offline using MATLAB, where gradient imperfections were corrected in an entirely data-driven fashion using AutoPSF\(^50\) which obviates the need for time-consuming calibration scans.

All multi-orientation data were processed offline using BET brain masking,\(^51\) FLIRT registration,\(^52\) Laplacian unwrapping,\(^53\) and SMV filtering.\(^54\) In order to increase the number of training samples for our deep learning reconstructions, data augmentation was performed on the QSMNet dataset (see Supporting Material Figure S1 for details). This increased the number of input/ground-truth pairs by a factor of five, resulting in a total of 45 datasets from nine subjects.

2.3 | Tikhonov regularization in NDI

The NDI optimization may start to fit noise and artifacts to further reduce the cost function if too many iterations are performed. As an alternative to early-stopping, we investigate the effect of Tikhonov regularization. It provides a framework to stabilize the solution of ill-conditioned linear equations and has been successfully applied to related problems such as parallel imaging.\(^55\) We examine the convergence of single- and multi-orientation NDI reconstructions with Tikhonov regularization (\(\lambda = 0.0\%, \lambda = 0.1\%, \lambda = 1.0\%\)) by computing the RMSE with respect to 5-direction NDI for different number of iterations throughout the optimization and across nine subjects (Supporting Material Figure S2). Additional assessments are provided in Supporting Material Figure S3, where we utilize SSIM as the error metric, compute RMSE with respect to 5-direction COSMOS and assess the convergence of NDI in a low signal-to-noise environment (SNR = 50).

2.4 | NDI vs. COSMOS, TKD, L2 and FANSI

Non-regularized NDI and COSMOS were compared for single- and multi-orientation reconstructions. To prevent over-fitting in NDI, the optimization was stopped after 450 iterations (see results from Figure 2). Moreover, single-orientation NDI with Tikhonov regularization (\(\lambda = 0.1\%, 450\) iterations, parameters obtained from Figure 2) was compared against TKD (\(k_{\text{trunc}} = 0.16\), L2 (\(\lambda = 3.1e-2\)),\(^38\) NMEDI (\(\lambda = 1400\), 4 iterations)\(^40\) and FANSI (\(r = 1e-5\), \(\mu_1 = 5e-2\), \(\mu_2 = 1.0\), 199 iterations)\(^41\) where the truncation and regularization parameters were tuned to minimize RMSE with respect to the 5-direction COSMOS data.

2.5 | Learning a variational regularizer for NDI (VaNDI)

We introduce VaNDI to further improve the reconstruction quality of NDI using a Variational Network (VN)\(^56\) which combines deep learning elements with the nonlinear QSM data model. This approach solves the following cost function

\[
\min_{\tilde{\chi}} \left\{ R(\tilde{\chi}) - \frac{1}{2} \| W (e^{\tilde{\phi}} - e^{\phi}) \|_2^2 \right\}
\]

where the regularization term \(R(\tilde{\chi}) = \sum_{i=1}^{N} \psi_i (K_i \tilde{\chi}, 1)\) is a generalization of the Fields of Experts model.\(^57\) Here, \(K_i\) denotes a linear operator that models a convolution with the kernel \(k_i\). Non-linear potential functions \(\psi_i\) are constructed from a weighted sum of scalar functions (radial basis functions). The gradient of this least-squares problem leads to an iterative gradient descent (GD) algorithm, which expands the NDI gradient update rule with learnt variational regularizers.
\[
\chi^{t+1} = \chi^t - \sum_{i=1}^{N_k} (K^i)^T \Psi^i_t (K^i \chi^t) - \lambda_t \nabla \chi^t W^T W \sin(D \chi^t - \dot{\varphi})
\]

In this equation, non-linear activation functions \( \Psi^i_t \) denote the first derivative of the potential functions \( \Psi^i_t \). Moreover, trainable parameters \( K^i_t \), \( \Psi^i_t \) and \( \lambda_t \) became time-dependent (compare\(^5\)), e.g. these functions are now specific to the \( t \)-th gradient descent step and \( i \)-th feature channel (if applicable). The time-varying regularization parameter \( \chi^t \) acts as a variable step size and dynamically weights the gradient update \( D^T W^T W \sin(D \chi^t - \dot{\varphi}) \) in each gradient descent step.

In our VaNDI implementation (Figure 1), 3D convolutions with kernel size 11x11x11 were used to generate \( N_k = 24 \) feature channels. Non-linear activation functions \( \Psi^i_t \) were parameterized with 21 radial basis functions. We used \( T = 5 \) gradient descent steps in our VaNDI implementation (note that 450 iterations were needed in NDI). Training was performed using the IPALM optimizer\(^5\) (batch size of one, and 800 epochs), where the \( \ell_2 \) loss between 1-direction VaNDI and 5-direction NDI was minimized. As benchmark of comparison, we also trained a 3D UNet\(^\text{59} \) with the same input/ground truth data as in VaNDI and the following network parameters: 4 max-pooling/deconvolution layers, 2 convolutional layers per max-pooling with a kernel size of 3x3x3, batch normalization and ReLU activation. For the training, the \( \ell_2 \) loss was minimized using ADAM\(^\text{60} \) (800 epochs). Both networks were trained on eight volunteers (five head orientations per subject, 40 datasets), and the ninth subject was used for testing.

2.6 | High-resolution NDI at 7 T with Wave-CAIPI encoding

We also assessed the performance of NDI on high-resolution GRE data from three head orientations acquired at 7 T. To facilitate such high-resolution acquisition in a reasonable timeframe, Wave-CAIPI encoding at \( R = 15 \)-fold acceleration was used. For benchmark of comparison, 3-orientation COSMOS was computed and compared to single- and multi-orientation NDI.

3 | RESULTS

In Figure 2, the effect of early-stopping and Tikhonov regularization is assessed. We computed the RMSE of 1–4-direction NDI with respect to 5-direction NDI for different Tikhonov regularization values. For both single- and multi-orientation reconstructions, NDI over-fitted the data when no regularization was used (\( \lambda = 0.0\% \), blue line). This is also demonstrated in the example reconstructions provided at the bottom of the figure. After 1500 iterations, NDI exhibited streaking artifacts, whereas after 50 iterations, NDI resulted in insufficient contrast with large gaps around the magic angle in k-space (c.f. Krylov sequence\(^23 \)). While the optimal stopping iteration with the smallest RMSE (blue dashed line) varied between the different reconstructions, a fixed small amount of Tikhonov regularization (\( \lambda = 0.1\% \), red line) prevented over-fitting for all observed numbers of directions and subjects. Moreover, this regularization parameter (\( \lambda = 0.1\% \)) seemed to provide robust reconstructions across the nine subjects (Supporting Material Figure S2) as well as in a low signal-to-noise environment (SNR = 50, see Supporting Material Figure S3).
Figure 3 compares NDI and COSMOS for single- and multi-orientation reconstructions. For small number of head orientations, COSMOS is subject to artifacts as \(\frac{\text{Pd} T_r \text{dr}}{C_{16}/C_{17}}\) is poorly conditioned. NDI could address this and improved the reconstruction quality dramatically even from a single-orientation input. At five head directions (angular range \([-18\; ; 25\])\), both techniques provided comparable image quality and contrast.

Figure 4 compares single-orientation NDI against parameter-optimized TKD, L2, NMEDI and FANSI. The regularization in L2, NMEDI and FANSI leads to blurring and over-smoothing when compared to NDI and TKD, which is also reflected in a larger data consistency error (see Table 1). In contrast, TKD provided sharper images but suffered from more streaking artifacts and contrast reduction, a consequence of the k-space underestimation around the magic angle (see k-space picture). A good trade-off between mitigation of artifacts and image sharpness was achieved by NDI, which also achieved the lowest data consistency error (magnitude weighting and complex formulation). The RMSE/SSIM metrics computed with respect to 5-direction COSMOS (Table 1) yielded overall comparable results for all algorithms under consideration and did not seem to reflect artifacts and loss of sharpness.

**Figure 2** Without regularization, NDI over-fits the data (\(\lambda=0.0\%\), blue line), if not stopped early (blue dashed line). This is also demonstrated in the example reconstructions provided at the bottom of the figure (insufficient contrast after 50 iterations, streaking artifacts from over-fitting after 1500 iterations). A small amount of Tikhonov regularization (\(\lambda=0.1\%\), red line) mitigated this issue robustly for all observed number of directions and subjects (see supporting material figure S2).

**Figure 3** Comparison of NDI vs. COSMOS for various numbers of head directions. NDI significantly reduced streaking artifacts and provided good results even at a single head orientation. In contrast, the 3-dir COSMOS reconstruction was still poorly conditioned (angles 0°, 17°, 25°) and resulted in streaking artifacts.
Figure 5 compares the results of 1-direction VaNDI (deep learning + nonlinear data-fidelity) and UNet. Both approaches overall improved the image quality when compared to single-direction NDI and resulted in better RMSE/SSIM (Table 2). Moreover, slight underestimation of the susceptibility signal in the single direction NDI input data was mitigated by both techniques as assessed quantitatively at the bottom of Figure 5. However, while UNet achieved good GM/WM contrast and overall crisp images, it introduced additional artifacts (marked with red arrows), an effect not observed in any of the VaNDI reconstructions. Moreover, both ML algorithms yielded increased data-consistency error (c.f. Table 2), which was likely caused by (i) susceptibility anisotropy in the 5-dir NDI reconstruction used in training as the ground truth, and (ii) sacrificing data consistency to reduce regularization cost. Note that 5-dir NDI also had larger data consistency error than 1-dir NDI and served as the ground truth in VaNDI and UNet, again attributable to susceptibility anisotropy.

The performances of NDI and COSMOS were also evaluated at 7 T using high-resolution data (0.5 mm isotropic) acquired with a WaveCAIPI GRE sequence (Figure 6). While 3-direction COSMOS resulted in poor image quality with streaking artifacts (max. angle of head rotation was 13°), 1-direction NDI provided better reconstructions. Further improvement was achieved using 3-direction NDI, where small anatomical features such as blood vessels and U-fibers (zoom-in) were more conspicuous.
DISCUSSION

Our data demonstrate high-quality QSM reconstructions from an arbitrary number of head orientations and suggest a robust and simple dipole inversion technique.

NDI does not use complicated regularizers (no spatial gradient penalty, TV, etc.) but relies on the inherent regularization effect introduced by magnitude weighting and nonlinear formulation. In addition, either early-stopping or a small amount of Tikhonov regularization was required to

**TABLE 2**  RMSE/SSIM and data consistency error are provided for 1-dir NDI, VaNDI, UNet and 5-dir NDI. Moreover, the reconstruction times are reported

|                         | 1-dir NDI | VaNDI   | UNet    | 5-dir NDI |
|-------------------------|-----------|---------|---------|-----------|
| RMSE w.r.t. 5-dir NDI   | 0.712     | 0.575   | 0.578   | -         |
| SSIM w.r.t. 5-dir NDI   | 0.845     | 0.927   | 0.910   | -         |
| Consistency $\|\text{Me}^{\text{GT}} - \text{Me}^{\text{N}}\|_2$ | 0.077     | 0.097   | 0.099   | 0.118     |
| Reconstruction time [s] | 133       | 4 (+133)| 4 (+133)| 288       |

**FIGURE 5** VaNDI and UNet further improved the image quality of single-direction NDI by reducing artifacts and improving the image contrast. Also, slight underestimation of the susceptibility signal in 1-dir NDI was mitigated by both ML techniques as evidenced by point cloud measurements and linear fits. However, while UNet better preserved the image sharpness, it introduced additional artifacts (red arrows) which were not observed in any of the VaNDI reconstructions. Moreover, imperfect flow compensation as well as registration and phase processing artifacts may have contributed in vessels (blue arrows) looking slightly different in single and multi-direction reconstructions where data from multiple head orientations was combined.
improve the convergence and prevent artifacts from over-fitting. A suitable Tikhonov regularization parameter was empirically determined, and robustness among different subjects, number of head orientations and signal-to-noise environments was observed. In practice, this may enable high-quality NDI reconstructions with small computational footprint and without the need to fine-tune regularization parameter(s) manually (pre-determined regularization).

In our experimental validation, NDI outperformed COSMOS for small number of head orientations (1–3) where the COSMOS reconstruction is poorly conditioned and numerically unstable. Using larger number of head orientations than three, NDI and COSMOS provided comparable image quality and contrast. Moreover, we demonstrated that NDI matches the RMSE/SSIM metrics from NMEDI and FANSI without the vulnerability of over-smoothing the images. Due to the small amount of regularization applied in NDI and TKD, these techniques also had lower data consistency error than L2, NMEDI and FANSI, which led to sharper images. With respect to reconstruction time, NDI yielded comparable results as FANSI (~2 min), however, we anticipate further speed-up using more advanced optimization techniques (e.g. non-linear conjugate gradient with backtracking line search\(^6\)). This should result in faster convergence and reduced reconstruction time.

We also investigated a novel deep learning approach to further refine the image quality by expanding NDI to admit variational regularizers learned from training data. Our VaNDI technique was compared to a UNet architecture, where comparable RMSE/SSIM was observed. However, while UNet achieved good contrast and overall crisp images it was more susceptible to artifacts when compared to our VaNDI approach and we have the following explanation for this: our UNet had a larger receptive field due to the higher number of hidden

---

**FIGURE 6**  High-resolution QSM data were generated from multi-orientation GRE scans at 0.5 mm isotropic resolution and R = 15x acceleration using Wave-CAIPI encoding. 3-direction COSMOS resulted in streaking artifacts, which was much improved using NDI even at a single orientation. The zoom-ins reveal fine anatomical features such as the U-fibers which are best seen in the 3-direction NDI reconstruction. However, despite extensive shimming, small artifacts were observed in regions of poor B0 homogeneity at the edge of the brain mask (see red arrows), which seemed to increase in strength when data from multiple head orientations were combined (both for NDI and COSMOS).
layers and the multi-scale resolution approach (pooling layers). While this property provided more representative power it also increased the risk of "hallucination", as observed in some of our reconstructions. In contrast, the receptive field of VaNDI was much smaller, yielding more controlled and localized regularization. In addition, VaNDI promotes solutions in agreement with the QSM model as the NDI gradient update rule was included in the network. In our experiments, these constraints seem to produce more robust and artifact-free reconstructions.

Ultimately, NDI was applied to 7 T where GRE data were acquired at 0.5 mm isotropic resolution with Wave-CAIPI encoding \((R = 15\times)\) and a custom tight-fitted coil to achieve high-quality imaging. This, however, limited the achievable head orientations to shallow angles which created a difficult dataset for QSM reconstructions, resulting in streaking artifacts in the COSMOS technique. In contrast, the inherent regularization of NDI enabled much better quality even at a single orientation, which further improved with more directions revealing fine anatomical structures such as iron content in the U-fibers.\(^{62,63}\) In contrast to previous publications,\(^{37}\) we also corrected for imperfections of the Wave gradients in an entirely data-driven fashion (using AutoPSF\(^{50}\)) without the necessity for time-consuming calibration scans. This enabled high-resolution QSM to become feasible in a much shorter acquisition/reconstruction time and should help pave the way for more frequent usage in the neuroscientific research community.

In conclusion, we developed a simple gradient descent optimizer to perform robust QSM with pre-determined regularization. We then combined NDI synergistically with deep learning where variational regularizers were learned from training data (VaNDI) to improve the image quality. Ultimately, we demonstrated the feasibility of high-resolution NDI at 7 T where Wave-CAIPI was utilized to facilitate highly accelerated acquisitions.

ACKNOWLEDGEMENTS

This work was supported in part by NIH research grants: R01EB020613, R01EB019437, P41EB015896, U01EB025162, the shared instrumentation grants: S10RR023401, S10RR019307, S10RR019254, S10RR023043, and NVIDIA GPU grants.

ORCID

Daniel Polak  https://orcid.org/0000-0001-9781-1528
Jongho Lee  https://orcid.org/0000-0002-9485-5434

REFERENCES

1. Liu T, Liu J, de Rochefort L, et al. Morphology enabled dipole inversion (MEDI) from a single-angle acquisition: comparison with COSMOS in human brain imaging. Magn Reson Med. 2011;65(3):777-783.

2. Langkammer C, Schweser F, Krebs N, et al. Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. Neuroimage. 2012;62(3):1593-1599.

3. Chen W, Zhu W, Kovanlikaya I, et al. Intracranial Calculifications and Hemorrhages: Characterization with Quantitative Susceptibility Mapping. Radiology. 2014;270(2):496–505.

4. Langkammer C, Kang Y, Zhang S, et al. Quantitative susceptibility mapping in multiple sclerosis. Radiology. 2013;267(2):551-559.

5. Langkammer C, Pirpamer L, Seiler S, et al. Quantitative susceptibility mapping in Parkinson’s disease. PLoS ONE. 2016;11(9):e0162460. 6. Acosta-Cabrero J, Williams GB, Cardenas-Blanco A, Arnold RJ, Lupson V, Nestor PJ. In vivo quantitative susceptibility mapping (QSM) in Alzheimer’s disease. PLoS ONE. 2013;8(11):e81093.

7. Fan AP, Bilic B, Gagnon L, et al. Quantitative oxygenation venography from MRI phase. Magn Reson Med. 2014;72(1):149-159.

8. Fan AP, Evans KC, Stout JN, Rosen BR, Adalsteinsson E. Regional quantification of cerebral venous oxygenation from MRI susceptibility during hypercapnia. Neuroimage. 2015;104:146-155.

9. Wang Y, Spincemaille P, Liu Z, et al. Clinical quantitative susceptibility mapping (QSM): biometal imaging and its emerging roles in patient care. J Magn Reson Imaging. 2017;46(4):951-971.

10. Hammond KE, Lupo JM, Xu D, et al. 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. Neuroimage. 2008;39(4):1682-1692.

11. Parker DL, Payne A, Todd N, Hadley JR. Phase reconstruction from multiple coil data using a virtual reference coil. Magn Reson Med. 2014;72(2):563-569.

12. Robinson S, Grabner G, Witoszynskij S, Trattnig S. Combining phase images from multi-channel RF coils using 3D phase offset maps derived from a dual-echo scan. Magn Reson Med. 2011;65(6):1638-1648.

13. Robinson SD, Dymerska B, Bogner W, et al. Combining phase images from array coils using a short echo time reference scan (COMPOSER). Magn Reson Med. 2017;77(1):318-327.

14. Jenkinson M. Fast, automated, N-dimensional phase-unwrapping algorithm. Magn Reson Med. 2003;49(1):193-197.

15. Cusack R, Papadakis N. New robust 3-D phase unwrapping algorithms: application to magnetic field mapping and undistorting echoplanar images. Neuroimage. 2002;16(3 Pt 1):754-764.

16. Abdul-Rahman HS, Gdeisat MA, Burton DR, Lalor MJ, Lilley F, Moore CJ. Fast and robust three-dimensional best path phase unwrapping algorithm. Appl Optics. 2007;46(26):6623.

17. Witoszynskij S, Rauscher A, Reichenbach JR, Barth M. Phase unwrapping of MR images using ΦUN - a fast and robust region growing algorithm. Med Image Anal. 2009;13(2):257-268.

18. Schofield MA, Zhu Y. Fast phase unwrapping algorithm for interferometric applications. Opt Lett. 2003;28(14):1194.
59. Ronneberger O, Fischer P, Brox T. “U-net: Convolutional networks for biomedical image segmentation,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2015, vol. 9351, pp. 234–241.

60. Kingma DP, Ba JL. “Adam: A method for stochastic optimization,” in 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, San Diego, USA, 2015.

61. Lustig M, Donoho D, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn Reson Med. 2007;58(6):1182-1195.

62. Deistung A, Schäfer A, Schweser F, Biedermann U, Turner R, Reichenbach JR. Toward in vivo histology: a comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and R2*-imaging at ultra-high magnetic field strength. Neuroimage. 2013;65:299-314.

63. Khabipova D, Wiaux Y, Gruetter R, Marques JP. A modulated closed form solution for quantitative susceptibility mapping - a thorough evaluation and comparison to iterative methods based on edge prior knowledge. Neuroimage. 2015;107:163-174.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Polak D, Chatnuntawech I, Yoon J, et al. Nonlinear dipole inversion (NDI) enables robust quantitative susceptibility mapping (QSM). NMR in Biomedicine. 2020;33:e4271. https://doi.org/10.1002/nbm.4271

APPENDIX

Analytical gradient derivation for NDI optimization

NDI is based on nonlinear-MEDI, where additional regularization terms were removed ($\alpha=0$). This enables an analytical gradient derivation as shown in previous work.40

NDI aims to minimize the following cost function

$$f(\vec{\chi}) = \| W(e^{iD\vec{\chi}} - e^{i\vec{\phi}}) \|^2$$

which can be rewritten as

$$= (e^{iD\vec{\chi}} - e^{i\vec{\phi}})^T W^T W (e^{iD\vec{\chi}} - e^{i\vec{\phi}})$$

Since $D\vec{\chi}$ is constrained to real values, we can use $(D\vec{\chi})^H = (D\vec{\chi})^T$

$$= (e^{-i(D\vec{\chi})^T} W^T W e^{iD\vec{\chi}} - e^{-i(D\vec{\chi})^T} W^T W e^{i\vec{\phi}} - e^{-iD\vec{\chi}^T} W^T W e^{i\vec{\chi}} + e^{-i\vec{\phi}^T} W^T W e^{i\chi})$$

Rewriting matrix multiplications using index notation yields

$$= \sum_c w^2_{cc} \left( 2 - e^{i(\sum_a \alpha_{ax} - \phi_a)} - e^{-i(\sum_a \alpha_{ax} - \phi_a)} \right)$$

Applying a trigonometric relation further simplifies the term

$$= 2 \sum_c w^2_{cc} \left( 1 - \cos \left( \sum_a \alpha_{ax} - \phi_a \right) \right)$$
Differentiating $f(\chi)$ with respect to $\chi_n$ yields

$$\frac{\partial f}{\partial \chi_n} = 2 \sum_c w^2_c D_{cn} \cdot \sin \left( \sum_b D_{bn} \frac{\phi_c}{\phi_b} \right)$$

Ultimately, we return to matrix notation and obtain

$$\nabla f(\vec{\chi}) = 2 D^T W^T W \sin \left( D \vec{\chi} - \vec{\phi} \right)$$