Supporting Information

Here we describe the Levenberg-Marquardt algorithm used for solving the blockwise curve-fitting problem in SUPER T1 mapping. The cost function to minimize is

$$\sum_{m=0}^{M-1}\sum_{l=0}^{L-1}\sum_{n=0}^{\frac{N}{R}-1}\left|\mathrm{y}_{lmn}-\mathbf{W}_l\mathbf{S}_{m\bar{n}}\left(\mathbf{A}_{\bar{n}}-\mathbf{B}_{\bar{n}}\exp\left(-\frac{t_l}{\mathbf{T1}_{\bar{n}}}\right)\right)\right|^2 \tag{S1}$$

Following our previous work in T2* mapping (1,2), we rewrite the cost function by change of variables

$$\sum_{m=0}^{M-1}\sum_{l=0}^{L-1}\sum_{n=0}^{\frac{N}{R}-1}|\mathrm{y}_{lmn}-\mathbf{W}_l\mathbf{S}_{m\bar{n}}(\mathbf{A}_{\bar{n}}-\mathbf{B}_{\bar{n}}\exp(\mathbf{Z}_{\bar{n}}t_l))|^2 \tag{S2}$$

where $\mathbf{Z}_{\bar{n}}=-1/\mathbf{T1}_{\bar{n}}$. Such a formulation provided an easier way to calculate the gradient. Notice that y_{lmn}, $\mathbf{A}_{\bar{n}}$, and $\mathbf{B}_{\bar{n}}$ are all complex-valued, thus a complex-valued gradient should be derived for optimization of the cost function. Alternatively, one can derive a separate gradient for the real and imaginary part of the variables (3). This transforms the cost function to

$$\sum_{m=0}^{M-1}\sum_{l=0}^{L-1}\sum_{n=0}^{\frac{N}{R}-1}\left|\mathrm{y}_{lmn}-\mathbf{W}_l\mathbf{S}_{m\bar{n}}\left(\left(\mathbf{A}_{\bar{n}}^{\mathrm{R}}+\iota\mathbf{A}_{\bar{n}}^{\mathrm{I}}\right)-\left(\mathbf{B}_{\bar{n}}^{\mathrm{R}}+\iota\mathbf{B}_{\bar{n}}^{\mathrm{I}}\right)\exp\left(\left(\mathbf{Z}_{\bar{n}}^{\mathrm{R}}+\iota\mathbf{Z}_{\bar{n}}^{\mathrm{I}}\right)t_l\right)\right)\right|^2 \tag{S3}$$

where $\mathbf{A}_{\bar{n}}^{\mathrm{R}}$, $\mathbf{B}_{\bar{n}}^{\mathrm{R}}$, and $\mathbf{Z}_{\bar{n}}^{\mathrm{R}}$ are the real part of the variable and $\mathbf{A}_{\bar{n}}^{\mathrm{I}}$, $\mathbf{B}_{\bar{n}}^{\mathrm{I}}$, and $\mathbf{Z}_{\bar{n}}^{\mathrm{I}}$ are the imaginary part of the variable. The separation of real and imaginary parts is aimed to facilitate separate bounding of each part in the optimization to avoid local minima. These boundaries include restricting $\mathbf{Z}_{\bar{n}}^{\mathrm{R}}\in[-0.01\ \mathrm{ms}^{-1},-0.0002\ \mathrm{ms}^{-1}]$, so that $\mathbf{T1}_{\bar{n}}\in[100\ \mathrm{ms},5000\ \mathrm{ms}]$, and $\mathbf{Z}_{\bar{n}}^{\mathrm{I}}=0$, since there should not be any phase variation in T1 relaxation signal. Furthermore, a non-zero significant phase term in the exponential term may cancel the modulation frequency from shift

undersampling and promote ill-conditioning and local minima. In this work, these boundaries were enforced by introducing regularization terms to the cost function S3:

$$\sum_{m=0}^{M-1}\sum_{l=0}^{L-1}\sum_{n=0}^{\frac{N}{R}-1}\left(\left|y_{lmn}-\mathbf{W}_l\mathbf{S}_{m\bar{n}}\left(\left(\mathbf{A}_{\bar{n}}^{\mathrm{R}}+\iota\mathbf{A}_{\bar{n}}^{\mathrm{I}}\right)-\left(\mathbf{B}_{\bar{n}}^{\mathrm{R}}+\iota\mathbf{B}_{\bar{n}}^{\mathrm{I}}\right)\exp\left(\left(\mathbf{Z}_{\bar{n}}^{\mathrm{R}}+\iota\mathbf{Z}_{\bar{n}}^{\mathrm{I}}\right)t_l\right)\right)\right|^2\right.$$
$$\left.+\Psi\left(\mathbf{Z}_{\bar{n}}^{\mathrm{R}},\lambda_1,\lambda_2,\mathrm{LB},\mathrm{HB}\right)+\lambda_3\left\|\mathbf{Z}_{\bar{n}}^{\mathrm{I}}\right\|^2\right) \tag{S4}$$

where λ_1, λ_2, and λ_3 are weights of each regularization, LB and HB are the lower and higher bound for $\mathbf{Z}_{\bar{n}}^{\mathrm{R}}$, respectively, and $\Psi\left(\mathbf{Z}_{\bar{n}}^{\mathrm{R}},\lambda_1,\lambda_2,\mathrm{LB},\mathrm{HB}\right)$ is a piecewise-defined, first-order differential function:

$$\Psi\left(\mathbf{Z}_{\bar{n}}^{\mathrm{R}},\lambda_1,\lambda_2,\mathrm{LB},\mathrm{HB}\right)=\begin{cases}\lambda_1\left(\mathbf{Z}_{\bar{n}}^{\mathrm{R}}-\mathrm{LB}\right)^2; & \text{if } \mathbf{Z}_{\bar{n}}^{\mathrm{R}}<\mathrm{LB}\\ 0; & \text{if } \mathrm{LB}\le\mathbf{Z}_{\bar{n}}^{\mathrm{R}}\le\mathrm{HB}\\ \lambda_2\left(\mathbf{Z}_{\bar{n}}^{\mathrm{R}}-\mathrm{HB}\right)^2; & \text{if } \mathbf{Z}_{\bar{n}}^{\mathrm{R}}>\mathrm{HB}\end{cases} \tag{S5}$$

Function S5 penalizes any $\mathbf{Z}_{\bar{n}}^{\mathrm{R}}$ that is lower than LB or higher than HB with a quadratic cost function. Notice these regularization terms are zero-valued when $\mathbf{Z}_{\bar{n}}^{\mathrm{R}}$ and $\mathbf{Z}_{\bar{n}}^{\mathrm{I}}$ stay within the bound. Thus, these regularization terms do not introduce any bias to the solution, since the physiological range is specified to ensure the solution belongs to this range. In our experience, we found that when λ_1, λ_2, and λ_3 are sufficiently large, the variables always converge to the solution. In our implementation, we used $\lambda_1 = 2\times10^5, \lambda_2 = 5\times10^9,$ and $\lambda_3 = 10^7$.

The cost function S4 is minimized with the Levenberg-Marquardt algorithm. A pseudocode for the algorithm can be found in (4). The initialization for the three parameters, A, B, and T1 were 1, 2, and 1000 in our implementation and yielded satisfactory performance. The maximal number of iterations was 300, and the algorithm was stopped when reaching the maximal number of

iterations or the norm of the gradient divided by the residual of the cost function was less than 1.0×10^{-6}.

Reference

1. Hu C, Reeves S, Peters DC, Twieg D. An Efficient Reconstruction Algorithm Based on the Alternating Direction Method of Multipliers for Joint Estimation of R2* and Off-Resonance in fMRI. IEEE transactions on medical imaging 2017;36(6):1326-1336.
2. Hu C, Reeves SJ. Trust Region Methods for the Estimation of a Complex Exponential Decay Model in MRI With a Single-Shot or Multi-Shot Trajectory. IEEE Transactions on Image Processing 2015;24(11):3694-3706.
3. Olafsson V, Fessler JA, Noll DC. Spatial resolution analysis of iterative image reconstruction with separate regularization of real and imaginary parts. In: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006. . IEEE. p 5-8.
4. Madsen K, Nielsen HB, Tingleff O. Methods for non-linear least squares problems. 1999.

Supporting Information Table S1. Reconstruction time per voxel for each reconstruction method in each experiment

Method	Acceleration Rate	Simulations		Phantom T1 mapping		Brain T1 mapping		MOLLI	
		Data Complexity	*Recon. time per voxel (ms)*	*Data Complexity*	*Recon. time per voxel (ms)*	*Data Complexity*	*Recon. time per voxel (ms)*	*Data Complexity*	*Recon. time per voxel (ms)*
Voxelwise curve-fitting	R=1	L=6; M=1*	0.6	L=7; M=1	0.6	L=8; M=1	0.7±0.1	L=11; M=1	1.0±0.1
SUPER	R=2	L=6; M=6	1.2	L=7; M=12	4.5	L=8; M=4	4.6±0.2	L=11; M=15	NA
MARTINI	R=2	L=6; M=6	22	L=7; M=12	174	L=8; M=4	79±3	L=11; M=15	NA
SUPER-SENSE	R=4	L=6; M=6	2.5	L=7; M=12	6.3	L=8; M=4	4.0±0.1	L=11; M=15	8.7±0.7
GRAPPATINI	R=4	L=6; M=6	28	L=7; M=12	153	L=8; M=4	72±3	L=11; M=15	NA

*: L is the number of TIs and M is the number of coils. For voxelwise curve-fitting, M=1 because the multi-channel image data is firstly combined before the parametric mapping reconstruction.

Supporting Information Table S2. Speed-up of SUPER to MARTINI and SUPER-SENSE to GRAPPATINI

Speed-up	Simulations	Phantom T1 mapping	Brain T1 mapping
SUPER/MARTINI	18	39	17±1
SUPER-SENSE/GRAPPATINI	11	24	18±1