Combined Approximate Message Passing for Total Variation Minimization and Randomly Translated Wavelet Denoising - Improved Compressed Sensing for Diffusion Spectrum Imaging

Jonathan I. Sperl[1], Ek T. Tan[2], Marion I. Menzel[1], Kedar Khare[2], Kevin F. King[3], Christopher J. Hardy[2], and Luca Marinelli[2]
[1]*GE Global Research, Garching n. Munich, BY, Germany,* [2]*GE Global Research, Niskayuna, NY, United States,* [3]*GE Healthcare, Waukesha, WI, United States*

Introduction: Diffusion spectrum imaging (DSI) [1], in which the diffusion encoding space (*q-space*) is Nyquist sampled, allows for resolving crossing fiber tracts in the brain or characterizing diffusional kurtosis [2]. To overcome its long acquisition times, q-space can be randomly undersampled and reconstructed using compressed sensing (CS) [3]. However, the iterative CS algorithms result in long computation times. This work presents several improvements to CS-DSI, namely the combination of the approximate message passing (AMP) [4] and Nesterov updates (NU) [5], the application of AMP to total variation (TV) minimization, and random translations (RT) [6] for wavelet transform (WT) based CS. Furthermore, all methods can be combined yielding superior convergence properties.

Theory: Given the undersampled q-space signal y, CS-DSI computes the data x in the reciprocal *r-space* by solving

$$\min_x \|Ax - y\|_2 + \lambda\|\Psi x\|_1 \qquad (1)$$

with $A=MF$ (undersampling operator and Fourier transform) and Ψ a sparsifying transform (TV or WT). Eq. 1 can be solved using iterative shrinkage algorithms (ISA) [7] (*cf.* Tab. 1). Ψ is incorporated via the denoising function η (*cf.* Tab. 2).

Recent modifications to ISA are NU and AMP. As shown in Tab. 1 *right,* both methods can be combined. The original AMP weight α includes the mean of the derivative of η [4] (*cf.* Tab. 3) with δ the ratio of the dimensions of domain and range of A, respectively. For soft thresholding (ST) [7] and the most simple case of Ψ = id, this term reduces to the fraction of values outside the interval $[-\sigma,\sigma]$. This approach can be transferred to general ST methods and also to nonlinear methods like TV (*cf.* Tab. 3 right).

The discrete WT is not shift invariant. However, RT [6] can be used, i.e. the WT Ψ in η_{ST} is replaced by ΨS_r (and the inverse Ψ^{-1} by $S_{-r}\Psi^{-1}$, respectively) whereas $S_r x$ denotes the circulant shift of x by r. In every iteration, r is chosen randomly, which yields approximate shift invariance. A combined approach (CMB=TV+WT+RT+AMP) applies $\eta_{ST,t}$ and $\eta_{TV,t}$ subsequently, whereas the AMP weight a is computed as the mean of $\alpha_{ST,t}$ and $\alpha_{TV,t}$.

standard algorithms			proposed
ISA [7]	ISA+NU [5]	ISA+AMP [4]	ISA+NU+AMP
$z^t = y - Ax^t$	$u^t = a_t x^t + b_t x^{t-1}$		$u^t = a_t x^t + b_t x^{t-1}$
	$z^t = y - Au^t$	$z^t = y - Ax^t + \alpha_t z^{t-1}$	$z^t = y - Au^t + \alpha_t z^{t-1}$
$w^t = A^H z^t + x^t$	$w^t = A^H z^t + u^t$	$w^t = A^H z^t + x^t$	$w^t = A^H z^t + u^t$
$x^{t+1} = \eta_t(w^t)$	$x^{t+1} = \eta_t(w^t)$	$x^{t+1} = \eta_t(w^t)$	$x^{t+1} = \eta_t(w^t)$

Tab. 1: ISA-Algorithms at iteration t.

soft thresholding (ST) [7] (Ψ e.g. WT)	TV [8]
$\eta_{ST,t}(w) = \Psi^{-1} T_{\sigma_t}(\Psi w)$	$\eta_{TV,t}(w) = w - \sigma_t \nabla \mathrm{TV}(w)$
$T_\sigma(s) = \begin{cases} s - \mathrm{sgn}(s)\sigma & \text{if } \lvert s\rvert \geq \sigma \\ 0 & \text{else.} \end{cases}$	$\nabla \mathrm{TV}(w) = -\mathrm{div}\left(\frac{\nabla w}{\lVert \nabla w \rVert}\right)$

Tab. 2: Denoising functions η.

original [4]	simplified: fraction of values above threshhold	
	ST	TV
$\alpha_{ST,t} = \delta\langle \eta_t'(w_t)\rangle$	$\alpha_{ST,t} = \delta \frac{\#\{\lvert \Psi w_t\rvert \geq \sigma\}}{\dim(\Psi w_t)}$	$\alpha_{TV,t} = \delta \frac{\#\{\lvert \Psi w_t\rvert \geq \lvert \sigma_t \nabla \mathrm{TV}(w_t)\rvert\}}{\dim(w_t)}$

Tab. 3: AMP weights α. (<·>: mean, #{·}: number of elements in set {·}, dim(x): dimension of x.)

Methods: Fiber simulations were performed using a Gaussian Mixture Model (2 fibers, fractional anisotropy FA=0.85, 70° crossing angle) on a 17^3 cube. Complex Gaussian noise was added (level 3%). The q-space data was randomly undersampled with various acceleration factors R. CS reconstruction was performed using ISA+NU and various combinations of the proposed techniques. Two error metrics were analyzed: RMSE between reconstruction and noiseless ground truth, and the error in the estimated fiber directions using the orientation distribution functions. Averages were taken over 50 sampling/noise patterns.

DSI experiments on healthy volunteers were performed using a 3T GE MR750 clinical MR scanner (GE Healthcare, Milwaukee, WI, USA) and an 8 Channel Head Coil (*TE =141 ms, TR = 3 s, 128x128,* b_{max} *= 10,000 s/mm*^{2}). The data were artificially undersampled (R = 4, 6) and CS reconstructed using ISA+NU and either TV only or CMB with 10, 15 or 20 iterations. Finally, kurtosis maps were analyzed by fitting the diffusion and kurtosis tensors.

Results: The simulation results in Fig. 1 show that WT only performed worst in terms of both error metrics. The application of RT substantially improved the WT result, while the application of AMP had slightly improved convergence over the TV only result. However, the combined method CMB was superior to all the other methods. In addition, the accuracy from using CMB at a higher acceleration factor may be equivalent to that from the TV only method at a lower acceleration factor, for example in the R = 6 vs. R = 4 comparison.

Fig. 2 shows kurtosis maps of the brain data. While the TV based reconstruction for R=4 appears still noisy and depicts the structures in the brain only weakly after 10 iterations, CMB already converged conveying the same information as the TV based method after 20 iterations. A similar improvement is visible when going to a higher acceleration factor R=6 and comparing both methods after 15 iterations.

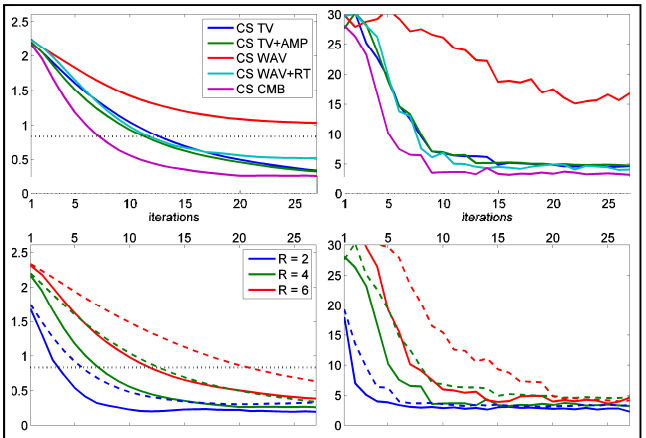

Fig. 1: Errors in fiber simulations. Left: logarithm of RMSE of q-space data, right: Error (in °) of estimated direction of one of the two fibers. Top: comparison of different CS recons for R=4. Bottom: comparison for different R (dashed line: CS TV, solid line: CMB). Black dotted line: error of fully sampled noisy data (no CS recon).

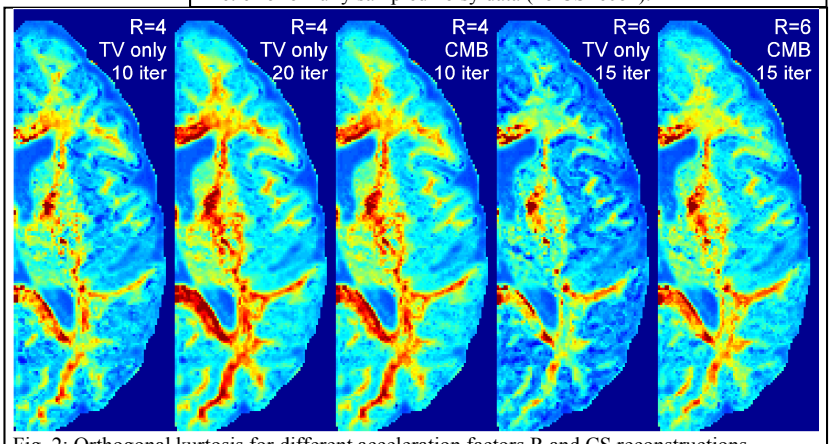

Fig. 2: Orthogonal kurtosis for different acceleration factors R and CS reconstructions.

Discussion: The proposed AMP and RT modifications yielded improved convergence properties when applied to the standard CS-DSI reconstructions using TV and wavelets, respectively. Moreover, combining all methods in a single reconstruction allows both shift invariant wavelet denoising and total variation minimization to be optimally used, yielding improved accuracy and faster convergence.

Acknowledgement: The authors would like to thank Prof. M. Lustig for helpful discussions on random translations.

References:

[1] Wedeen *et al.*, MRM., 2005 [2] Jensen *et al.*, NMR Biomed, 2010
[3] Menzel *et al.*, MRM 2011 [4] Donoho *et al.*, PNAS 2009
[5] Khare *et al.*, MRM, *under review*
[6] Figueiredo *et al.*, IEEE TIP 2003
[7] Daubechies *et al.*, Commun. Pure Appl. Math. 2004
[8] Sidky *et al.*, Phys. Med. Biol. 2008