

MIT Open Access Articles

Joint multi#contrast variational network reconstruction (jVN) with application to rapid 2D and 3D imaging

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

As Published: 10.1002/MRM.28219

Publisher: Wiley


Persistent URL: <https://hdl.handle.net/1721.1/135944>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of use: Creative Commons Attribution 4.0 International license



Joint multi-contrast variational network reconstruction (jVN) with application to rapid 2D and 3D imaging

Daniel Polak^{1,2,3}  | Stephen Cauley^{2,4,5} | Berkin Bilgic^{2,4,5} | Enhao Gong⁶ | Peter Bachert^{1,7} | Elfar Adalsteinsson⁸ | Kawin Setsompop^{2,4,5}

¹Department of Physics and Astronomy, Heidelberg University, Heidelberg, Germany

²Department of Radiology, A. A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA

³Siemens Healthcare GmbH, Erlangen, Germany

⁴Harvard Medical School, Boston, MA, USA

⁵Harvard-MIT Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA, USA

⁶Subtle Medical Inc, Menlo Park, CA, USA

⁷Medical Physics in Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

⁸Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA

Correspondence

Daniel Polak, A. A. Martinos Center for Biomedical Imaging, 149 13th Street, Charlestown, MA 02129, USA.
Email: daniel.m.polak@gmail.com

Funding information

National Institute of Health, Grant/Award Number: P41EB015896, R01EB019437, R01EB020613, R01MH116173, S10RR019254, S10RR019307, S10RR023043, S10RR023401 and U01EB025162; Nvidia, Grant/Award Number: GPU grant

Purpose: To improve the image quality of highly accelerated multi-channel MRI data by learning a joint variational network that reconstructs multiple clinical contrasts jointly.

Methods: Data from our multi-contrast acquisition were embedded into the variational network architecture where shared anatomical information is exchanged by mixing the input contrasts. Complementary k-space sampling across imaging contrasts and Bunch-Phase/Wave-Encoding were used for data acquisition to improve the reconstruction at high accelerations. At 3T, our joint variational network approach across T1w, T2w and T2-FLAIR-weighted brain scans was tested for retrospective under-sampling at $R = 6$ (2D) and $R = 4 \times 4$ (3D) acceleration. Prospective acceleration was also performed for 3D data where the combined acquisition time for whole brain coverage at 1 mm isotropic resolution across three contrasts was less than 3 min.

Results: Across all test datasets, our joint multi-contrast network better preserved fine anatomical details with reduced image-blurring when compared to the corresponding single-contrast reconstructions. Improvement in image quality was also obtained through complementary k-space sampling and Bunch-Phase/Wave-Encoding where the synergistic combination yielded the overall best performance as evidenced by exemplary slices and quantitative error metrics.

Conclusion: By leveraging shared anatomical structures across the jointly reconstructed scans, our joint multi-contrast approach learnt more efficient regularizers, which helped to retain natural image appearance and avoid over-smoothing. When

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Magnetic Resonance in Medicine* published by Wiley Periodicals, Inc. on behalf of International Society for Magnetic Resonance in Medicine

synergistically combined with advanced encoding techniques, the performance was further improved, enabling up to $R = 16$ -fold acceleration with good image quality. This should help pave the way to very rapid high-resolution brain exams.

KEYWORDS

deep learning, Joint multi-contrast reconstruction, parallel imaging, Wave-CAIPI

1 | INTRODUCTION

Fast imaging techniques have been widely adopted into clinical practice to speed up MRI scans and, thus, help improve patient throughput, reduce the sensitivity to involuntary patient motion,¹ improve patient compliance, and potentially obviate the need for sedation in pediatric patients.² However, conventional parallel imaging (PI) algorithms (eg *SENSE*,³ *GRAPPA*,⁴ etc) are constrained to moderate acceleration rates, R , (eg typically $R = 3$ for 2D and $R = 2 \times 2$ for 3D) to avoid structural artifacts and large noise amplification. To enable higher accelerations with improved image quality, advanced encoding and reconstruction techniques have been proposed.

Among these techniques, *2D-CAIPIRINHA*⁵ is applicable to volumetric 3D acquisitions and employs a staggered k_y - k_z under-sampling pattern to create controlled aliasing in the phase (y) and partition (z) encoding plane, which increases the distance between the aliasing voxels and enables better utilization of coil sensitivity information in the reconstruction. *Wave-CAIPI*⁶ adopts this scheme and combines it with *Bunch Phase Encoding (BPE)*⁶ by playing additional sinusoidal gradients on both the G_y and G_z gradients with a quarter-cycle phase shift during the readout. This enables controlled aliasing along all three spatial dimensions, including the readout axis (x), which significantly reduces artifacts and g-factor noise amplification when compared to *2D-CAIPIRINHA*. At 3T, the *Wave-CAIPI* technology was demonstrated to provide up to 9-fold acceleration for 3D sequences⁷⁻⁹ with comparable diagnostic quality as *GRAPPA* at the lower acceleration rate of $R = 4$. Moreover, *Wave-CAIPI* was employed in Simultaneous Multi-Slice (*SMS*) sequences, where there is no \sqrt{R} -SNR penalty from multiband (MB) acceleration, allowing up to $R = 12$ -fold effective acceleration.¹⁰ However, in general the efficiency of this technique is significantly reduced when applied to 2D sequences (without *SMS*) where controlled aliasing is limited to the x - y domain (cf. *BPE*). Besides artifacts also SNR can be a challenge at very high acceleration (both for 2D and 3D) due to the inherent \sqrt{R} -noise penalty and may necessitate going to higher magnetic field strength (as in *Wave-GRAPPA*¹¹) or using frameworks like *Compressed Sensing (CS)*¹² and *LORAKS*,¹³ which have also been synergistically combined with *Wave-CAIPI*.^{14,15}

However, for techniques like *CS* to work robustly, several pre-requisites need to be fulfilled. Aliasing artifacts must be incoherent which is commonly achieved by non-Cartesian or random under-sampling, but since most clinical sequences employ Cartesian sub-sampling, incoherence is in practice limited to dynamic and 3D imaging sequences but remains a challenge for 2D acquisitions. Moreover, *CS* requires the existence of a representation in which the reconstructed images become sparse. Commonly used transformations are wavelet,¹² total-variation (TV),¹⁶ and total generalized variation (TGV),¹⁷ which in combination with the ℓ_1 norm achieve at least approximate sparsity. However, the use of the ℓ_1 norm entails iterative optimization algorithms which are often computationally demanding and yield longer reconstruction times. Also, the choice of the regularization parameter(s) is critical to balance between over-smoothing and residual artifacts.

Recent developments in deep learning have the potential to lift some of these barriers. On highly accelerated data, neural networks have outperformed existing techniques both in terms of image quality, artifact reduction as well as reconstruction time. The algorithms proposed in Hyun et al and Polak et al^{18,19} operate on coil-combined images and were trained to un-alias zero-padded reconstructions or enhance the image quality of conventional methods such as *SENSE*, *GRAPPA* or *CS*, etc. Moreover, further improvement was demonstrated by reconstructing multiple clinical contrasts jointly. This idea was previously investigated for *PI-CS* reconstructions where additional sparsity constraints along the contrast dimension were used,²⁰⁻²³ and this concept has now also been applied to deep learning.²⁴ By exploiting the redundancy across the jointly reconstructed contrasts, these techniques enable better image quality than single-contrast methods. However, the pixel-wise loss used in these approaches requires the multi-contrast data to be spatially registered, which may pose a challenge for clinical routine. A recent work²⁵ discovered the relevance of this issue and proposed a conditional GAN with cyclic consistency loss²⁶ to jointly reconstruct unregistered multi-contrast data.

Several groups have demonstrated the benefits of incorporating the multi-channel MRI data into the deep learning reconstruction. *RAKI*²⁷ is a k -space based technique where a convolutional neural network (CNN) is trained

to synthesize non-acquired lines in k-space. When compared to *GRAPPA*, which is a linear interpolation method, this non-linear extension yields improved noise resilience at high acceleration. Also, *RAKI* may be favorable as the training is performed solely on the subject specific ACS data; hence, large amounts of training data are not required. *AUTOMAP*²⁸ takes this one step further by learning the entire transformation from under-sampled multi-channel k-space data to the final image without ever explicitly using the Fourier transformation. This may present a flexible alternative for the reconstruction of non-Cartesian k-space trajectories where the exact inverse transform may not exist.

Inspired by traditional iterative techniques for inverse problems, several approaches²⁹⁻³¹ have posed the MRI image reconstruction as an unrolled gradient descent optimization where the physics model is embedded in the reconstruction and regularizers/priors are learnt from training data. This formulation can be understood as a generalization of *CS* where neural networks are utilized instead of hand-crafted domain transformations (such as wavelet or TV). With this framework many existing physics- and *CS*-based techniques have been outperformed while enabling much shorter reconstruction times.²⁹ In a recent work³² such a network was also utilized to reconstruct a highly accelerated *Wave* acquisition where imperfections of the sinusoidal *Wave* gradient trajectory were automatically estimated by the network without additional time-consuming optimizations (eg AutoPSF³³).

In this contribution, we augment the unrolled gradient descent optimization in the variational network (*VN*) architecture²⁹ to jointly reconstruct multiple clinical contrasts (T1w, T2w, T2-FLAIR) from accelerated MRI acquisitions. By taking advantage of the shared anatomical information across the jointly reconstructed scans, our joint variational network (*jVN*) approach learns more efficient regularizers which improved the image quality when compared to single-contrast *VN* reconstructions. Moreover, we investigated how complementary k-space sampling across imaging contrasts and advanced acquisition techniques such as *BPE* and *Wave*-encoding can be utilized to further boost the reconstruction performance. We validated these techniques both on 2D and 3D data and ultimately demonstrate the feasibility of obtaining T1w, T2w and T2-FLAIR contrasts at 1 mm isotropic resolution with $R = 16$ -fold acceleration in less than three minutes of scan time.

2 | METHODS

2.1 | Network architecture of *jVN*

This work is based on the variational network (*VN*) architecture,²⁹ which aims to solve the PI problem as an unrolled gradient descent (GD) optimization (Figure 1), where each step contains a traditional data-fidelity operation as well as a learnt regularizer. In the *VN* architecture the regularization

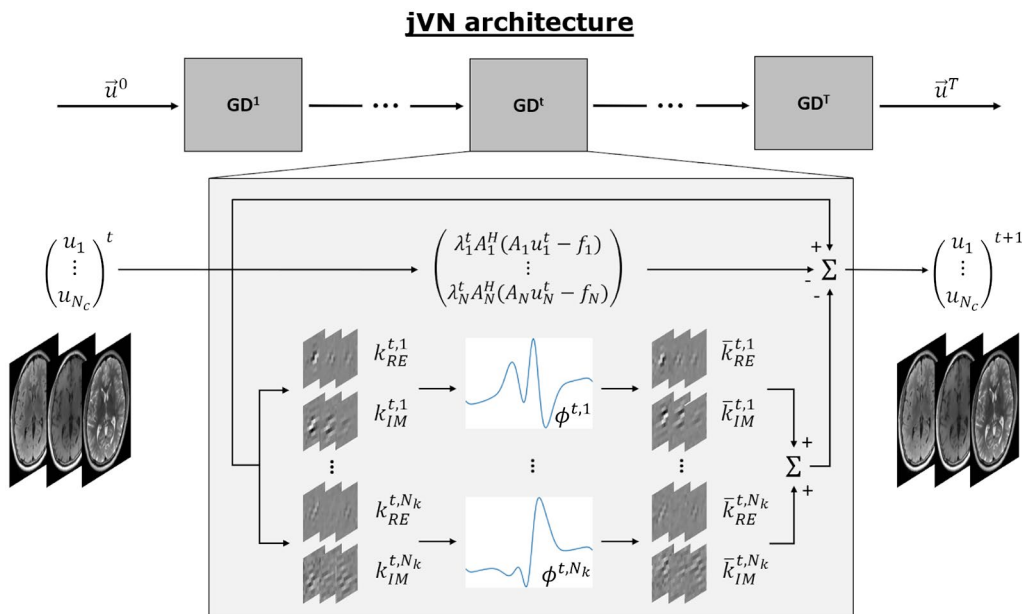


FIGURE 1 *jVN* is based on the variational network architecture²⁹ and poses the image reconstruction as an unrolled gradient descent (GD) optimization. Each gradient descent step GD^t contains a convolutional filter k^t which mixes the different input contrasts \vec{u}^t and creates N_k feature channels. Non-linear activation ϕ^t and the transposed filter \bar{k}^t reduce the N_k feature channels to the number of input contrasts N_c . Data-fidelity is computed individually for each contrast, where each forward model matrix A_c contains a contrast-specific under-sampling mask that can vary between contrasts to enable complementary k-space sampling (compare Figure 2). For *BPE*/*Wave* acquisitions, A_c additionally contains the *Wave* point-spread-function (Psf) to account for the voxel spreading along the readout direction

term was based on the Fields of Expert model,³⁴ a generalization of TV, where linear convolutional filter and non-linear potential functions are learnt during the training. Our joint variational network (*jVN*) augments this technique by reconstructing multiple clinical contrasts simultaneously. This is achieved by stacking eg T1w, T2w and T2-FLAIR-weighted images along the channel dimension of the network. Starting from an initial reconstruction \vec{u}^0 , each gradient descent step mixes the N_c input contrasts by convolving them with the filter kernels k^t resulting in N_k feature channels. Next, learned activation functions ϕ^t and the transposed filter \vec{k}^t are applied to reduce the number of feature channels to the number of input contrasts. Note, that the transposed filter kernels \vec{k}^t are applied when calculating the gradient update rule for the Fields of Expert model (see Hammernik et al²⁹ for details). Moreover, a data-fidelity term $A_c^H(A_c u_c^t - f_c)$ is computed individually for each contrast c (no mixing between scans), where A_c denotes the parallel imaging encoding matrix, u_c^t the current image estimate and f_c the undersampled multi-channel scanner data. The data-fidelity is weighted by a trained regularization parameter λ_c^t and subtracted from \vec{u}^t during each iteration.

As explored by previous contributions,³⁵ the image quality of multi-contrast reconstructions can be improved using complementary k-space under-sampling, eg by varying the acceleration factor across the input contrasts and/or including a contrast-dependent shift in the k-space sampling. This results in different aliasing and image artifacts in the initial *SENSE*-based reconstructions of the different image contrasts, which can be leveraged in multi-contrast reconstructions. In this work, we chose to keep the acceleration factor fixed but shift the uniform under-sampling pattern (see Figure 2) for each contrast. Particularly in 3D, this approach simplifies the data handling as the coupled voxel locations are identical across all contrasts; note, that we use *jVNc* to refer to joint variational network reconstructions with complementary k-space sampling.

Moreover, we utilized *BPE* to improve the quality of our 2D scans and *Wave*-encoding for our 3D acquisitions. Since

these techniques couple the readout dimension (x) into the PI problem, the encoding matrix is no longer separable along this dimension and reconstructing a full dataset at once may be intractable on state-of-the-art GPUs (especially for high-resolution 3D scans). To mitigate this issue, we constrained our acquisitions to uniformly under-sampled k-space masks with fixed acceleration. This allowed the PI reconstruction to be split into smaller sub-problems of collapsing voxels in image space. In our implementation, we divided the reconstructed 3D volume into stacks of coupled 2D slices, eg, at $R = 4 \times 4$ acceleration, $R_z = 4$ coupled 2D slices were fed to the network and reconstructed simultaneously. We modified the network's forward model operator $A_c = \sum_{z_j} \sum_{y_i} \mathcal{F}_x^{-1} \text{Psf} \mathcal{F}_x C p_c$ ⁷ to account for these changes where the index c denotes the contrast dependency and \mathcal{F}_x the Fourier transformation along x . The A_c operator first applies a linear phase ramp p_c to reflect shifts in the uniform k-space sampling mask and then multiplies with the coil sensitivity C and *Wave* point-spread-function *Psf* in hybrid space $[k_x, y, z]$. Ultimately, A_c sums over the collapsing voxels y_i, z_j ($i \in [1 \dots R_y], j \in [1 \dots R_z]$) corresponding to the acceleration factors R_y and R_z .

Furthermore, the following modifications were implemented: As no internal autocalibration scan (ACS) data were used for any of the reconstructions (ie no fully sampled k-space center), we generated input images u^0 from an initial *SENSE*-based instead of a zero-padded reconstruction (in contrast to Hammernik et al²⁹). We empirically observed that this improved the image quality for all evaluated reconstructions. Moreover, we trained individual networks for every output contrast, which was found to provide overall better image quality than a single network. We implemented this in the training stage by extracting one image contrast from the vector \vec{u}^T (containing all jointly reconstructed contrasts) before minimizing the ℓ_2 -loss with respect to the corresponding ground truth data. In this way, the loss function only measured the fidelity in a single contrast instead of all the input images.

Complementary k-space undersampling

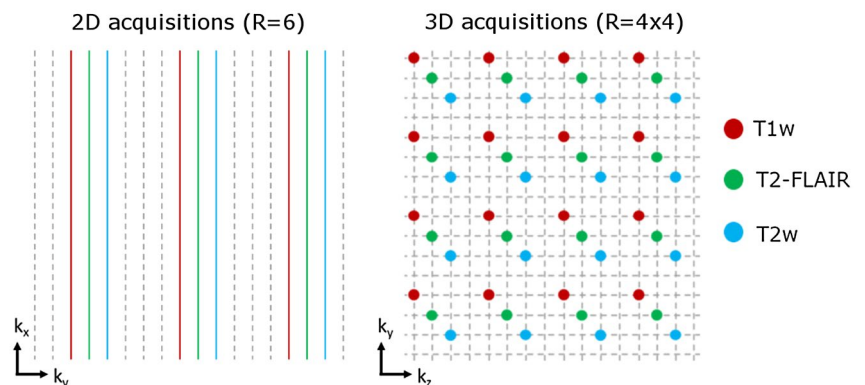


FIGURE 2 Joint multi-contrast reconstructions (*jVNc*) employ complementary k-space under-sampling by imposing a contrast-dependent shift on the uniform sub-sampling mask

2.2 | Data acquisition and pre-processing

2.2.1 | Retrospective acceleration

With IRB approval and informed consent, fully sampled training data were acquired on nine healthy subjects using two 3T scanners (MAGNETOM Prisma and Skyra, Siemens Healthcare, Erlangen Germany) and a product *SPACE* sequence (variable flip angle 3D turbo spin echo³⁶) with T1w, T2w and T2-FLAIR-weighted contrasts (field of view: $256 \times 256 \times 192 \text{ mm}^3$, resolution: $1 \times 1 \times 1 \text{ mm}^3$, orientation: sagittal, BW: 592Hz/px, product Siemens 32-channel head coil, more details are summarized in Supporting Information Table S1). Next, the acquired datasets were co-registered channel-by-channel using FSL FLIRT,³⁷ to mitigate any inter-scan motion between the acquisitions of T1w, T2w and T2-FLAIR-weighted scans.

This multi-contrast data were utilized for accelerated 2D and 3D imaging where a variety of sampling strategies were analyzed. For our 2D experiments, the fully sampled data were reformatted into axial orientation with 1 mm in plane resolution and 4 mm slice thickness (whole-brain coverage). For 3D, sagittal orientation and 1 mm isotropic resolution were retained. Next, the central 20×20 lines of k-space were extracted from the T2-FLAIR-weighted scan and coil sensitivity maps were computed using ESPIRiT³⁸; *BPE* (for 2D) and *Wave*-encoding (for 3D) were synthesized by convolving the fully sampled datasets with a point-spread-function (Psf) corresponding to the following *Wave* acquisition parameters: four sinusoidal *Wave* cycles per readout with 16 mT/m maximum gradient amplitude; two-fold readout oversampling (see Ref. [9] and Supporting Information Figure S1). We retrospectively undersampled the data at $R = 6$ for 2D and $R = 4 \times 4$ for 3D where in both cases no integrated ACS lines were kept. Each dataset was then reconstructed in MATLAB using *SENSE* with Tikhonov regularization. The corresponding regularization parameter was optimized to reduce NRMSE with respect to the fully sampled reference.

2.2.2 | Prospective acceleration

On three separate subjects, additional scans were performed to test prospective acceleration. For these acquisitions at $R = 4 \times 4$ (no integrated ACS lines), we used a prototype *Wave SPACE* sequence⁹ with complementary k-space sampling and the *Wave* acquisition parameters described above. The combined acquisition time for T1w, T2w and T2-FLAIR-weighted *SPACE* was $TA = 2:53$ min including a two second external *GRE* reference scan for the computation of coil sensitivity maps. Imperfections of the sinusoidal *Wave* gradient trajectory were estimated entirely data-driven using AutoPSF³³ (no additional calibration scans). The resulting *Wave* Psf was then

utilized in the initial *SENSE* and all variational network reconstructions. For benchmark of comparison, the same contrasts were also acquired without *Wave*-encoding at $R = 4 \times 4$ and $R = 2 \times 2$ and reconstructed using *SENSE*.

2.3 | Training and testing

To assess the benefit of reconstructing multiple contrasts jointly and/or utilizing complementary k-space sampling and *BPE/Wave*, separate networks were trained while the following parameters were held constant: $T = 10$ iterations, $N_k = 24$ real/imaginary filter pairs with filter size 11×11 , learned activations from 31 radial basis functions distributed equally between $[-1, 1]$. For each contrast, initial *SENSE* reconstructions u_c^0 and acquired k-space data f_c were individually pre-normalized by $u_c^0 = \frac{u_c^0}{\|u_c^0\|_2}$, $f_c = \frac{f_c}{\|f_c\|_2}$. All pre-processing steps were implemented in MATLAB, while training was performed in TensorFlow (Python) using a Nvidia GeForce GTX1080 GPU and the IPALM optimizer³⁹ (250 epochs). For our 2D scans, 1008 axial slices from seven subjects were used for training (batch size: 5), testing was performed on two subjects who were not used in the training.

We also characterized potential artifacts in the presence of inter-scan motion where the jointly reconstructed scans are not spatially registered. The performance of a preliminary motion correction technique was evaluated for these experiments. Fully sampled data were acquired on a separate subject with and without instructed inter-scan motion ($\Omega = 1.7^\circ$ in-plane rotation for T2w and $\Omega = -3.1^\circ$ for T1w; negligible translation and through-plane rotation). We reconstructed these test datasets using our *jVNC + BPE* network that was solely trained on registered images. Moreover, we evaluated a preliminary correction technique where registration of the image estimates \tilde{u}' was performed within each stage of the unrolled network. For this, additional translation and rotation operators were placed before the convolutional filter k' and corresponding inverse transformations after \tilde{k} to retain agreement with the acquired scanner data (note that translation and rotation operations were implemented in TensorFlow using an image transform function with bilinear interpolation). We tested this setup using the estimated motion parameters that were obtained by registering the initial *SENSE* reconstructions \tilde{u}^0 (MATLAB *imregister*) and compared the image quality and NRMSE to corresponding reconstructions without inter-scan motion.

For our 3D datasets at $R = 4 \times 4$, separate networks were trained with and without complementary k-space sampling and *Wave*-encoding using the same architecture as described for 2D. However, as each training sample now consisted of four sagittal slices ($R_z = 4$), the batch size was reduced to two to limit the required GPU memory. Overall, 336 training samples (1344 slices) from seven subjects were used for training, while testing was performed on 48 samples (192 slices) from

one subject. Moreover, VN and $jVNc + Wave$ were tested on prospective scans from three subjects at $R = 4 \times 4$ acceleration where Wave-encoding was used in the acquisition and was not synthesized.

3 | RESULTS

Figure 3 demonstrates the results for T2-FLAIR at $R = 6$ -fold acceleration. As shown, the encoding capability of $SENSE$ was insufficient at such high acceleration, causing large noise amplification and residual aliasing artifacts. However, the single-contrast VN mitigated most of these issues, but the artifact and noise reduction came at the cost of over-smoothing and loss of spatial resolution, as indicated by the zoom-in.

When reconstructing T1w, T2w and T2-FLAIR-weighted contrasts jointly ($jVNc$) or utilizing BPE ($VN + BPE$), the NRMSE was decreased (NRMSE was computed across 36 slices). Moreover, fine anatomical details were also better preserved as demonstrated by the improved conspicuity of a region of cerebrospinal fluid (CSF) in the posterior of the brain for $VN + BPE$ (red arrows). However, the overall best performance was achieved when $jVNc$ was synergistically combined with BPE . This is best seen in the zoom-in, where for example the gray-white matter boundary (blue arrows) or a small line of CSF (green arrows) became visible which were over-smoothed in all other reconstructions. These improvements are also reflected in better NRMSE, SSIM and PSNR which are provided in Table 1. Similar results were also obtained on the second test subject as demonstrated in

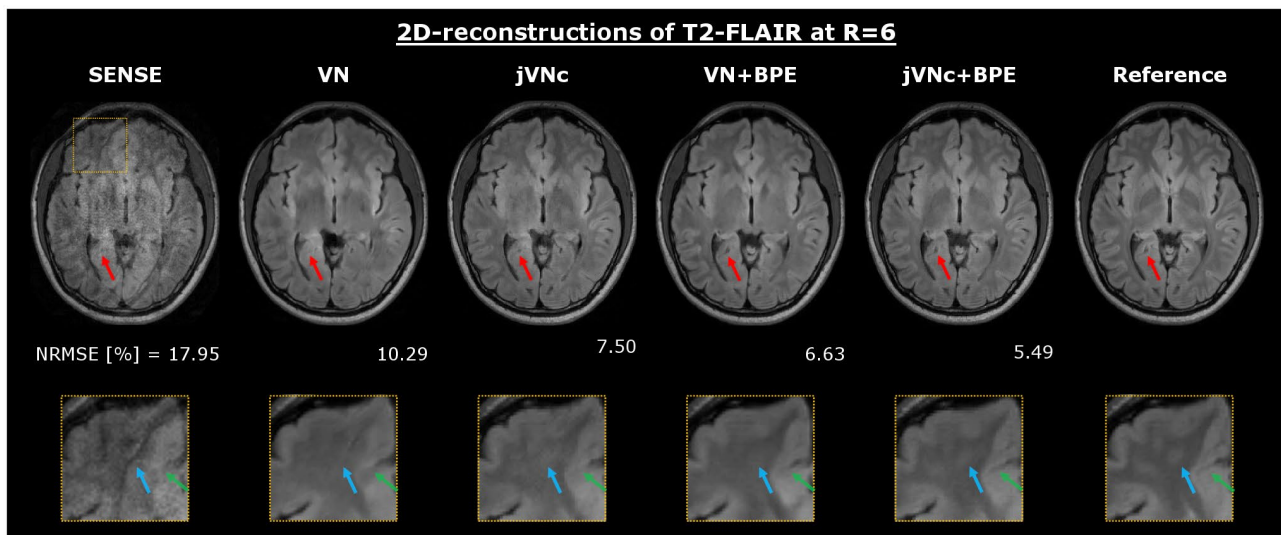


FIGURE 3 At $R = 6$ -fold acceleration, the $SENSE$ reconstruction of T2-FLAIR resulted in large noise amplification and aliasing artifacts, which were mostly mitigated using the single-contrast VN . However, by reconstructing T1w, T2w and T2-FLAIR-weighted contrasts jointly ($jVNc$) or utilizing BPE ($VN + BPE$), fine anatomical details were better preserved and the over-smoothing reduced. The overall best performance was achieved by $jVNc + BPE$ which is also reflected in the lowest NRMSE (computed across 36 slices)

TABLE 1 Quantitative metrics (NRMSE, SSIM, and PSNR) computed across 36 slices (test subject #1) are provided for T2-FLAIR, T2w and T1w reconstructions at $R = 6$ -fold acceleration

		<i>SENSE</i>	<i>VN</i>	<i>jVN</i>	<i>jVNc</i>	<i>VN + BPE</i>	<i>jVN + BPE</i>	<i>jVNc + BPE</i>
T2-FLAIR	NRMSE [%]	17.95	10.29	8.69	7.50	6.63	5.88	5.49
	SSIM	0.834	0.948	0.960	0.968	0.975	0.981	0.983
	PSNR	31.56	36.39	38.00	39.31	40.26	41.42	41.96
T2w	NRMSE [%]	18.25	10.51	8.55	7.14	6.08	5.59	5.25
	SSIM	0.855	0.952	0.968	0.975	0.983	0.986	0.986
	PSNR	32.62	37.22	39.00	40.56	41.96	42.68	43.24
T1w	NRMSE [%]	13.39	9.83	7.99	7.22	5.59	5.21	4.94
	SSIM	0.885	0.948	0.961	0.957	0.982	0.985	0.985
	PSNR	35.68	38.21	40.09	41.25	43.05	43.75	44.34

Note: Improvement over VN was achieved by either reconstructing all contrasts jointly (jVN), employing complementary under-sampling ($jVNc$) or utilizing BPE. The overall best results were obtained from the synergistic combination ($jVNc + BPE$) and are highlighted in bold.

Supporting Information Table S2. Moreover, a comparison to single- and multi-contrast *PI-CS* with corresponding error metrics is provided in the Appendix (Figure A1).

Figure 4 displays the results for T1w, T2w and T2-FLAIR at $R = 6$ -fold acceleration. Across all imaging contrasts, *jVNc + BPE* better retained the spatial resolution when compared to *VN* which is best seen in the anterior part of the brain (red arrows) where the gray-white matter boundary is over-smoothed. Moreover, the comparison demonstrates that the sequence-specific contrast was preserved without signal leaking from one scan to another. This is best seen in the center of the brain where green arrows mark a circular region of CSF (low signal intensity in T1w and T2-FLAIR, hyper-intense in T2w scan). Despite the non-linear mixing of all input contrasts in the convolutional filters of the joint variational network, no change in signal intensity was observed in any of the reconstructions,

while the conspicuity of this anatomical feature was much improved when compared to *VN* and *SENSE*. Moreover, *jVN* distinguished the structure of CSF from the structure of a blood vessel (blue arrows) even though the shape and size were similar (low-signal intensity of blood was retained for all contrasts).

In Figure 5, the effect of inter-scan motion in joint multi-contrast reconstructions was analyzed. Our *jVNc + BPE* network without motion correction resulted in residual aliasing artifacts (green arrows) when the input images were not aligned. Using our preliminary motion correction technique such artifacts were largely mitigated providing comparable image quality to that observed without inter-scan motion. This is also reflected in the quantitative NRMSE metric, where our correction technique enabled up to 19% reduction and only slightly higher NRMSE when compared to the no-motion reconstructions.

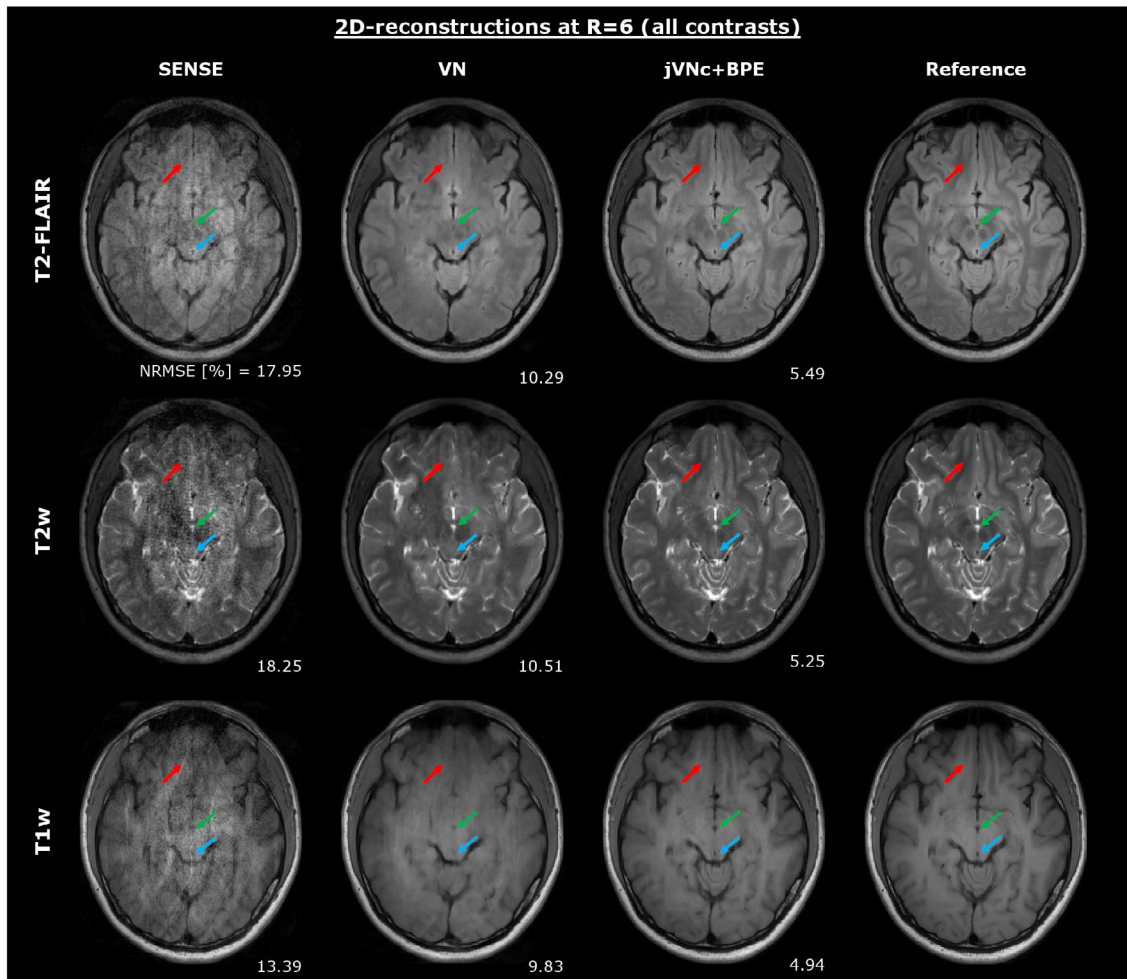


FIGURE 4 Throughout all contrasts, *jVNc + BPE* better preserved the spatial resolution (red arrows) and achieved lower NRMSE compared to *VN*. Moreover, the comparison demonstrates that *jVNc + BPE* retained the scan-specific contrast. Signal leakage from one contrast to another was not observed, as exemplarily demonstrated for a region of CSF (green arrows) which is hyper-intense in T2w, but dark in T1w and T2-FLAIR. Moreover, *jVN* distinguished the structure of CSF from the structure of a blood vessel (blue arrows) even though the shape and size were similar (low-signal intensity of blood was retained for all contrasts)

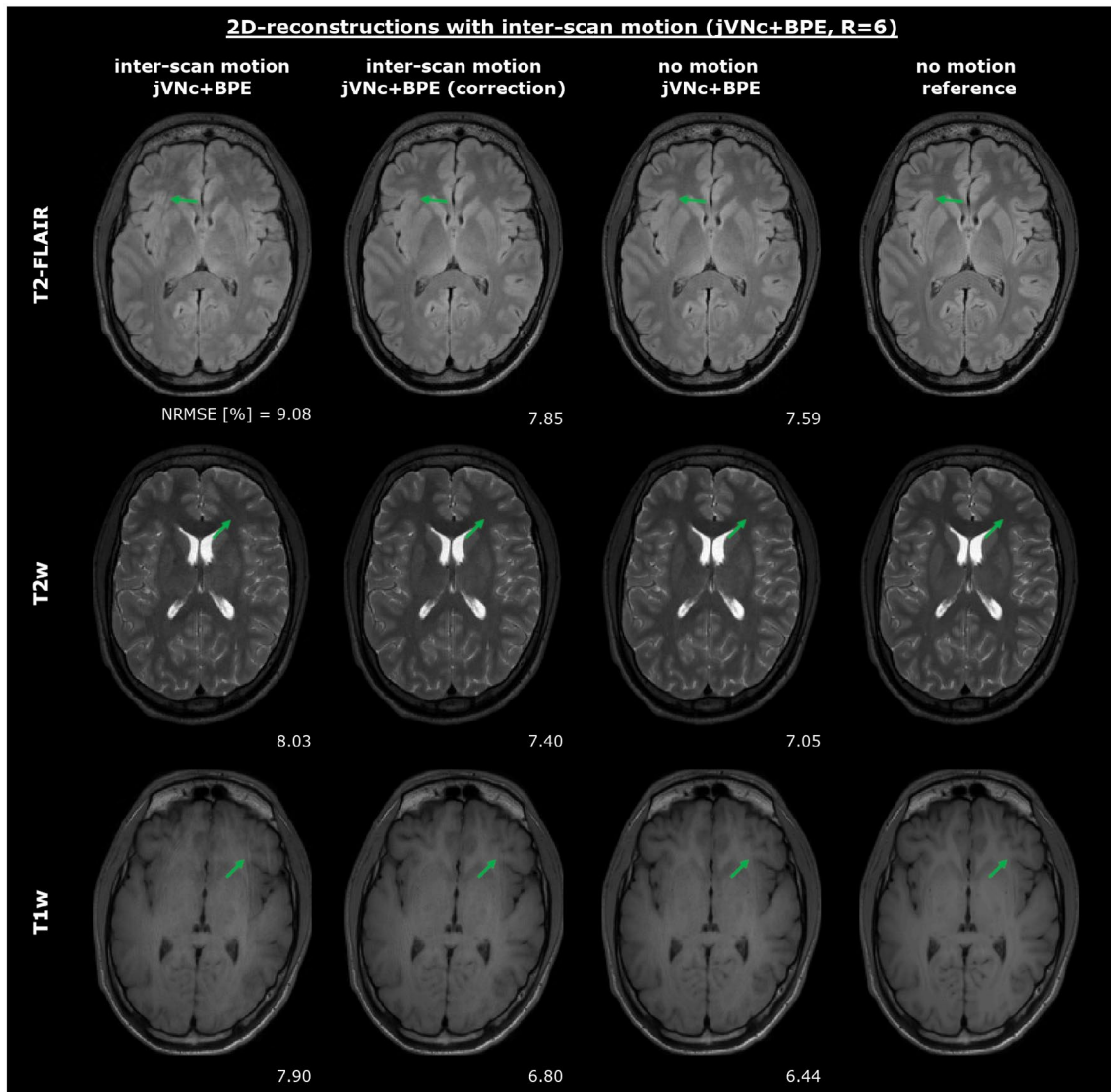


FIGURE 5 Inter-scan motion caused artifacts (green arrows) in our joint multi-contrast reconstructions ($jVNc + BPE$), which were significantly reduced using our motion mitigation technique. This enabled up to 19% reduction in NRMSE and comparable image quality as obtained without inter-scan motion

The results of the 3D reconstructions at $R = 4 \times 4$ acceleration are displayed in Figure 6. Again, the *SENSE* reconstruction suffered from \sqrt{R} -SNR loss and g-factor noise amplification. In contrast, the *VN* mitigated large noise enhancement, but the coronal reformats exhibit residual aliasing (green arrows) and striping artifacts (blue arrows) due to the ill-conditioning of the reconstruction which was performed sequentially across the aliasing coronal slice groups (convolutional filters in *VN* were applied to sagittal cuts). Moreover, the zoom-in reveals residual aliasing and loss of spatial resolution (red arrows) in regions of high g-factor where the encoding capability of the 32-channel head coil is limited. In contrast, $jVNc$ with complementary under-sampling helped to reduce some of these artifacts and improved NRMSE, but the striping artifacts were only mitigated when the PI problem

was better conditioned using *Wave*. The lowest NRMSE was obtained by $jVNc + Wave$, as demonstrated in the zoom-in, where the gray-white matter boundary was best preserved.

Finally, we tested our variational networks on prospectively accelerated data ($R = 4 \times 4$) from three test subjects acquired with and without *Wave*-encoding. The results are displayed in Figures 7 and 8, where a conventional acquisition at $R = 2 \times 2$ acceleration (no *Wave*-encoding) served as the reference. Both variational network reconstructions were able to preserve the sequence-specific contrast, however, $jVNc + Wave$ more efficiently removed aliasing artifacts (green arrows) and better preserved the spatial resolution (red arrows) across all test subjects. Nevertheless, at such high acceleration also $jVNc + Wave$ suffered from slight image blurring, eg in the cerebellum of T2-FLAIR.

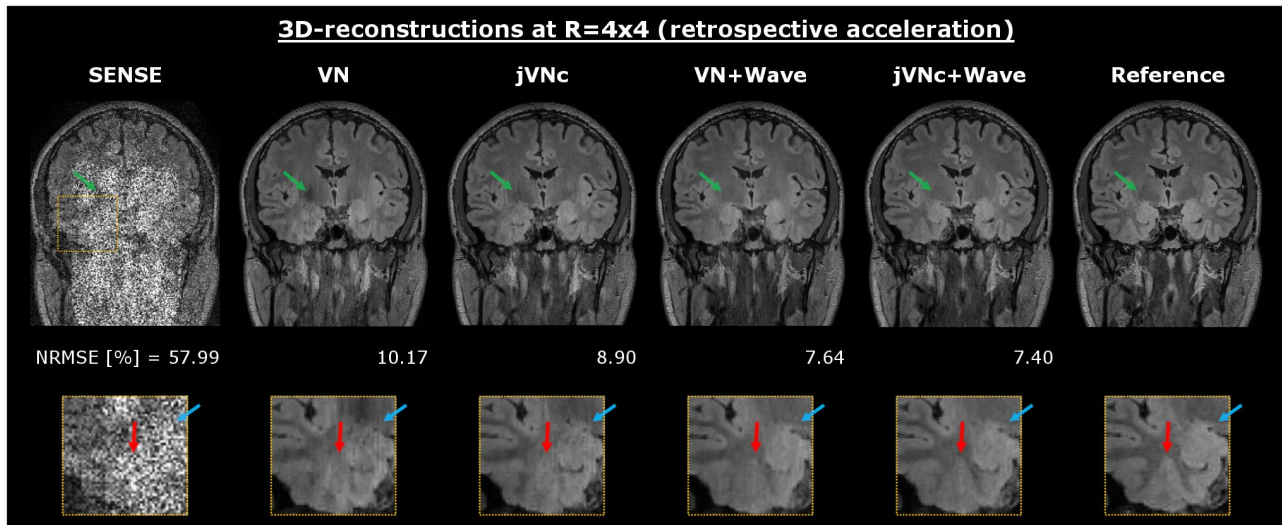


FIGURE 6 At $R = 4 \times 4$ acceleration, *VN* efficiently denoised the initial *SENSE* reconstruction but resulted in residual aliasing (green arrows), striping artifacts (blue arrows) and over-smoothing (red arrows). This was improved in multi-contrast *jVNc*, however, striping artifacts were only mitigated in the *Wave* reconstructions. The overall best performance was obtained by *jVNc + Wave*

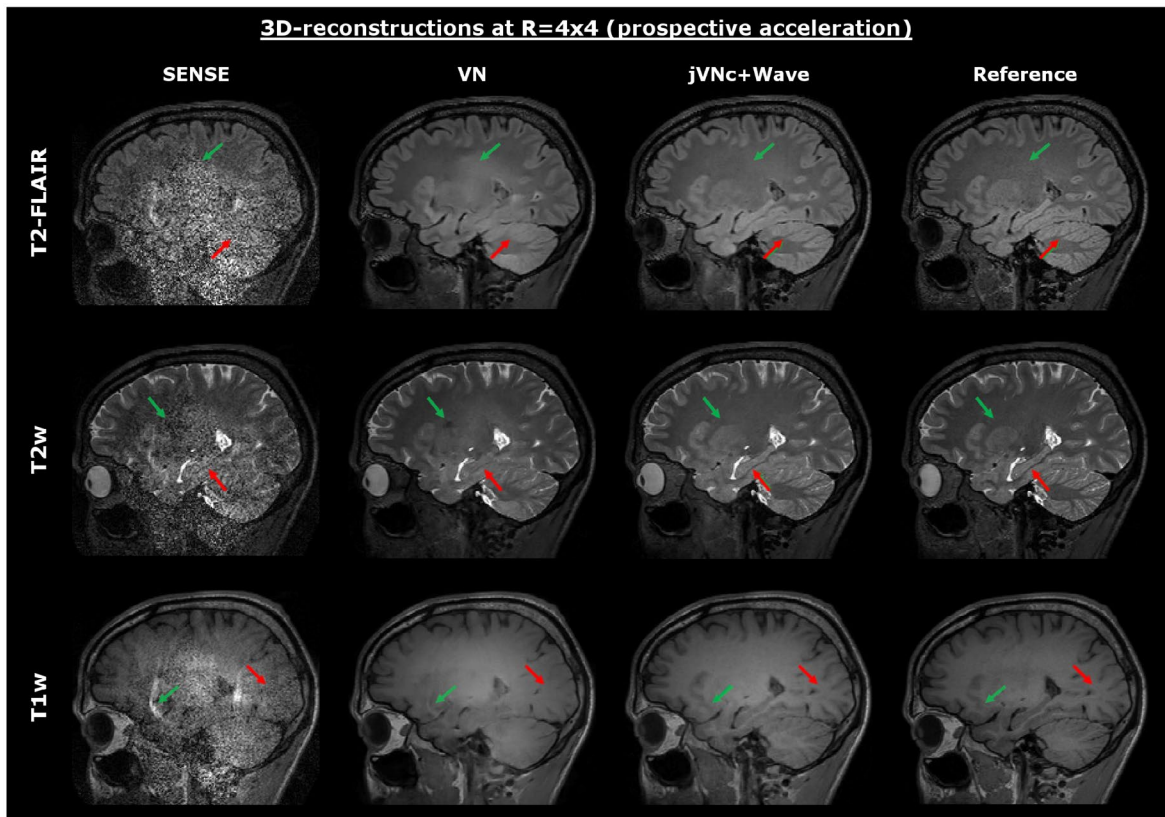


FIGURE 7 The variational networks were tested on prospectively accelerated data acquired at $R = 4 \times 4$ acceleration (combined TA = 2:53 min). The sequence specific contrast was retained in all scans, but *jVNc + Wave* better preserved fine anatomical details (red arrows) and exhibits fewer artifacts (green arrows) than *VN*. Nevertheless, at such high acceleration ($R = 16$) also the *jVNc + Wave* reconstructions resulted in small image blurring, for example in the cerebellum of T2-FLAIR

Moreover, the results demonstrate that our networks generalized to prospective acquisitions although the training data were under-sampled retrospectively and *Wave*-encoding was synthesized.

4 | DISCUSSION

In this contribution, we developed a framework to reconstruct data from multiple clinical imaging contrasts jointly

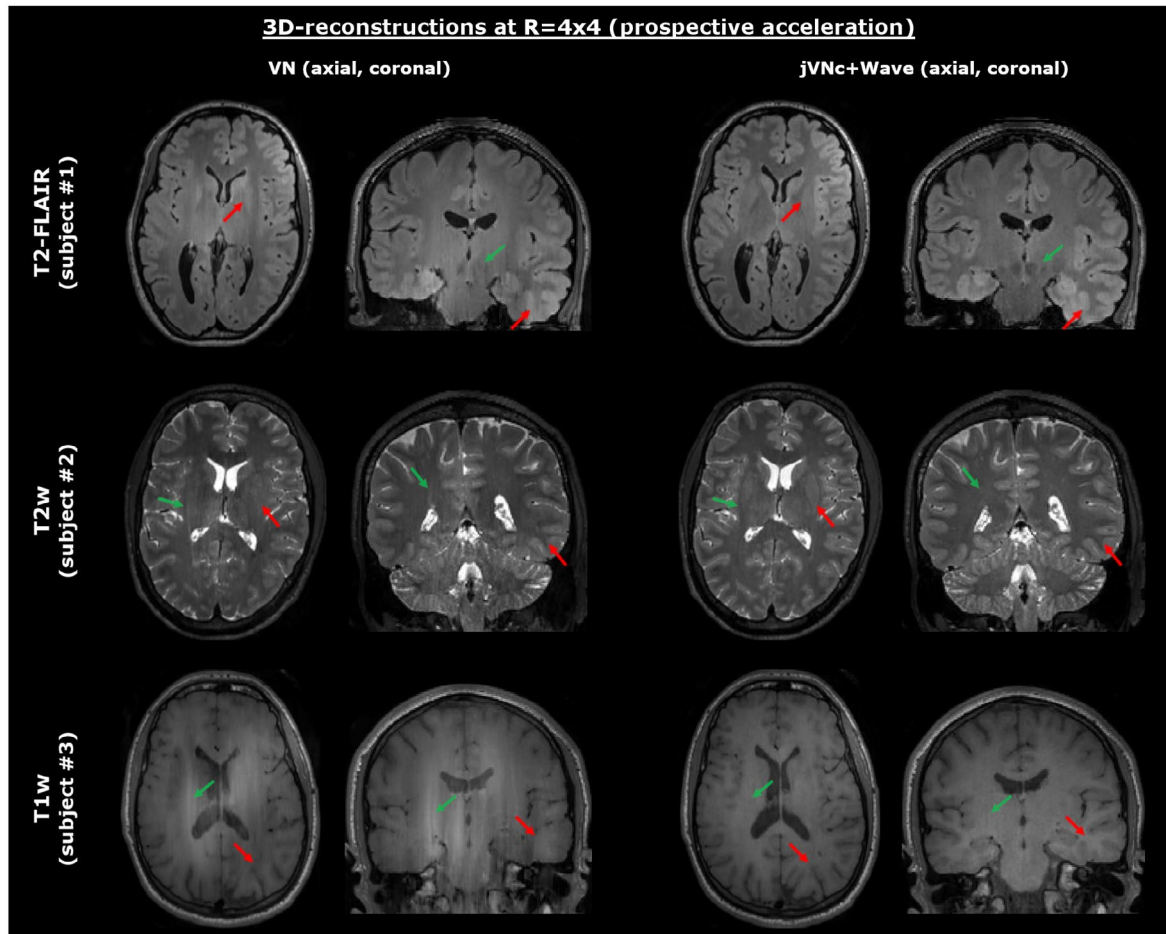


FIGURE 8 *VN* and *jVNc + Wave* were tested on three subjects at $R = 4 \times 4$ acceleration. Across all test cases, *jVNc + Wave* better mitigated aliasing artifacts (green arrows) and loss of spatial resolution (red arrows)

using the variational network architecture. By utilizing shared anatomical information across the imaging contrasts, *jVN* learned more efficient regularizers, which enables the reconstruction of highly under-sampled datasets with significantly reduced artifacts and image blurring. Moreover, we incorporated advanced encoding techniques in our acquisitions and demonstrated the benefit of complementary k-space under-sampling and *BPE/Wave*-encoding. This allowed T1w, T2w and T2-FLAIR-weighted scans to be acquired and jointly reconstructed at $R = 6$ -fold acceleration for 2D and up to $R = 16$ -fold acceleration for 3D (combined TA < 3 min), while retaining good image quality.

We quantitatively assessed the benefits from reconstructing multiple contrasts jointly and/or utilizing advanced encoding schemes such as *BPE* or *Wave* and showed that the synergistic combination yielded the overall best results. While the former technique allows the network to learn more efficient regularizers by leveraging shared anatomical structures across the jointly reconstructed contrasts, the latter improves the overall conditioning of the PI reconstruction by exploiting variations of the coil sensitivity also along the readout. It was observed that the combined

approach enabled higher improvement for 2D compared to 3D, where at $R = 6$ -fold acceleration the standard *SENSE* reconstructions resulted in large residual aliasing due to the insufficient encoding capability. In contrast, the $R = 16$ -fold accelerated *SENSE*-reconstructions for 3D were less effected by artifacts but dominated by the \sqrt{R} -SNR and g-factor noise penalties ($\sqrt{R} = 4$). This suggests that learning more efficient regularizers in *jVNc* mainly helps to resolve structural aliasing (as in our 2D scans) but is less beneficial in the presence of low SNR and few artifacts (such as in our 3D scans).

Our joint variational network was also compared to single- and multi-contrast *PI-CS* reconstructions (see Appendix). Here, additional regularization across the contrast dimension helped to preserve some fine anatomical details but only yielded slight image quality improvements. The learnt regularization in *jVN* better exploits shared anatomical structures and outperformed *PI-CS*. This may be a result of the vast contrast differences between T1w, T2w and T2-FLAIR-weighted scans where there is typically little to no linear dependence. Moreover, the joint *TGV* regularization restricted exchange of shared information to the spatial gradient domain, while

learnt filters in jVN operate directly on the multi-contrast image data and thus provide more degrees of freedom than traditional approaches.

We also assessed the performance of jVN in the presence of inter-scan motion, where the different clinical contrasts were not spatially aligned. Such motion may occur eg between pre- and post-contrast acquisitions, where there is typically a delay of several minutes. Our preliminary results on 2D data with inter-scan motion revealed that spatial misalignment in joint multi-contrast reconstructions may result in residual aliasing artifacts. This is intuitively clear, as the training was solely performed on registered data and the network learned to leverage this property as an additional prior in the reconstruction. However, such artifacts were almost entirely removed using our proposed motion mitigation technique which embeds translation and rotation operators in the network and utilizes the initial *SENSE* reconstructions to estimate the motion parameters. While good image quality was obtained on all test datasets, NRMSE was slightly higher when compared to the no-motion reconstructions. This was likely caused by small inaccuracies in the motion parameter estimation which was performed on the initial *SENSE* reconstructions \tilde{u}^0 . Better performance may be obtained when estimating these transformation variables on single-contrast *VN* reconstructions which typically provide higher image quality. Moreover, the bilinear interpolation used in all translation and rotation operations causes image smoothing which affects the data-fidelity term. However, we anticipate that more advanced interpolation techniques such as *spline* or *sinc* can reduce these effects at the cost of some increase in computation time. Also, it is important to note that rotations clip the corners of k-space and hence lead to some loss of spatial resolution. While not present in this work, further complications might arise from through-plane and intra-scan motion. In the latter case, prospective⁴⁰ or retrospective (data-driven) correction techniques^{41,42} could help to reduce associated artifacts.

Also, special attention was paid to potential artifacts caused by the mixing of clinical contrasts in the convolutional filters. While this was found to be beneficial for the reconstruction of highly accelerated datasets, it bears the risk of signal leaking from one scan to another, which could impede the clinical diagnosis. All test slices were carefully reviewed, however, such artifacts were never observed and we have the following explanation for this. In jVN multiple data-fidelity computations are embedded throughout the feed-forward path of the network which helps to hold the contrast-mixing in check. If signal leaked from one scan to another, the data inconsistency between the current estimate of the image and the acquired scanner data would increase. Also, due to the coupling in the PI problem, artifacts would not only remain at the location of origin but spread to all coupled voxel locations, a global

penalty which the filters in jVN efficiently learnt to avoid. Nevertheless, due to the small number of test subjects available in this study, further investigation with larger patient cohorts is necessary to confirm these initial findings especially in the presence of pathology.

In this work, a joint multi-contrast reconstruction was performed across T1w, T2w and T2-FLAIR-weighted scans which are commonly used in clinical brain exams. To ensure that jVN provided superior reconstruction quality than *VN* for every output contrast, individual joint networks were trained. This improved the ability of jVN to leverage shared anatomical structures, despite large contrast and SNR differences between the jointly reconstructed scans and enabled consistent image enhancement over *VN* for all output contrasts. Moreover, we anticipate that further acceleration feasibility can be achieved by increasing the number of clinical contrasts, eg by including T2*w (SWI)⁷ and/or post-contrast T1w⁸ and/or diffusion weighted imaging (DWI). This would not only provide improved regularization from increased anatomical information but also enable more efficient complementary k-space sampling and could pave the way for a very rapid multi-contrast brain exam^{9,43}. However, in such undertaking, it will be important to assess and potentially refine the jVN architecture to enable robust reconstructions across imaging contrasts with large background phase differences. In this work, all scans were acquired using a 3D TSE sequence, which resulted in the same image phase across all contrasts. However, phase variations may also arise from the coil sensitivity maps themselves which in the development phase of this work were calculated from the fully sampled k-space data of each imaging contrast individually and then included in the forward model of the reconstruction. It was observed that such phase differences can cause degradation in the reconstruction performance when compared to reconstructions that use the same set of coil sensitivities across all the contrasts. We anticipate that such behavior is specific to the *VN* architecture, where real and imaginary feature maps are summed after the convolutional filtering (k') and are not being kept as separate channels³². The integration of diffusion weighted imaging (DWI) into jVN may be challenging for several reasons: DWI is typically acquired at a smaller matrix size than most structural scans which will require interpolation to match the field of view and resolution of the remaining scans. Moreover, our current jVN implementation requires each contrast to be distortion free (similar spatial features), whereas DWI typically exhibits artifacts from several off-resonance effects, which need to be corrected. Susceptibility induced distortions can be reduced by increasing the acceleration factor. Moreover, modelling the B0 field in the MRI forward model was shown to mitigate distortions in multi-shot EPI with Blip Up-Down Acquisitions (BUDA)⁴⁴ and we

expect that necessary adaptations in the data term can be implemented into jVN . Residual artifacts from eddy currents and subject motion are often corrected as a post-processing step,⁴⁵ these corrections may be inserted into jVN before and after the convolutional filters. However, at high acceleration these transformation variables accounting for eddy currents and subject motion might be hard to estimate and may require repeated application of jVN to improve the DWI reconstruction while refining necessary parameters.

Variational networks were utilized in this work, as these architectures combine elements of deep learning with classical physics-based reconstructions. This allows advanced encoding techniques such as complementary k-space sampling and BPE/Wave-encoding to be integrated into the network and leveraged in the reconstruction. However, we expect that further image quality improvements can be obtained from deeper network architectures and/or expanding regularization in jVN to all three spatial dimensions. Currently, such implementations exceeded our GPU compute power and memory limitations but should be feasible on more advanced hardware. In this work, we employed 2D convolutional filters (regularization in x, y and contrast dimension), as 3D convolutions would have resulted in a 50-fold increase in GPU memory consumption (assuming that the entire 3D dataset was reconstructed at once). Instead we propose to expand our 2D technique by including surrounding slices in the reconstruction rather than the entire volume. This would allow regularization also in the partition-encoding direction using 3D convolutions but at much smaller computational cost and memory demand. Moreover, we expect that this strategy will be particularly suitable for 3D Wave acquisitions which are commonly SNR- and not encoding-limited. In such reconstructions the primary objective is the mitigation of noise for which the localized signal statistics is most relevant. Also, more sophisticated loss functions, such as SSIM,⁴⁶ and/or adversarial loss⁴⁷ can be utilized to improve the image sharpness and preservation of fine details.^{48,49} In contrast to ℓ_2 which measures pixel-wise differences, SSIM considers the signal intensity statistics in localized patches and hence provides a better representation of the human perceptual system.⁵⁰ Similarly, adversarial loss may help to reduce blurring and improve image sharpness especially in the presence of low SNR,⁴⁹ which would be beneficial for our T2-FLAIR reconstructions.

In conclusion, we demonstrated the benefit of reconstructing multiple clinical contrasts jointly and investigated how complementary under-sampling and BPE/Wave-encoding can be facilitated to improve the image quality. We carefully evaluated the performance of our networks both on 2D and 3D acquisitions, analyzed potential artifacts from inter-scan motion and finally demonstrated the feasibility of obtaining T1w, T2w and T2-FLAIR-weighted contrasts at high isotropic resolution in less than three minutes of scan time.

ACKNOWLEDGMENTS

This work was supported in part by NIH research grants: R01EB020613, R01EB019437, R01 MH116173, P41EB015896, U01EB025162, the shared instrumentation grants: S10RR023401, S10RR019307, S10RR019254, S10RR023043, and NVIDIA GPU grants.

CONFLICTS OF INTEREST

My Ph.D. salary is paid by Siemens Healthcare GmbH.

ORCID

Daniel Polak  <https://orcid.org/0000-0001-9781-1528>

REFERENCES

- Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: a complex problem with many partial solutions. *J Magn Reson Imaging*. 2015;42:887-901.
- Barton K, Nickerson JP, Higgins T, Williams RK. Pediatric anesthesia and neurotoxicity: what the radiologist needs to know. *Pediatr Radiol*. 2018;48:31-36.
- Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med*. 1999;42:952-962.
- Griswold MA, Jakob PM, Heidemann RM, et al. Generalized auto-calibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med*. 2002;47:1202-1210.
- Breuer FA, Blaimer M, Mueller MF, et al. Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA). *Magn Reson Med*. 2006;55:549-556.
- Moriguchi H, Duerk JL. Bunched phase encoding (BPE): a new fast data acquisition method in MRI. *Magn Reson Med*. 2006;55:633-648.
- Bilgic B, Gagoski BA, Cauley SF, et al. Wave-CAIPI for highly accelerated 3D imaging. *Magn Reson Med*. 2015;73:2152-2162.
- Polak D, Setsompop K, Cauley SF, et al. Wave-CAIPI for highly accelerated MP-RAGE imaging. *Magn Reson Med*. 2018;79:401-406.
- Polak D, Cauley S, Huang SY, et al. Highly-accelerated volumetric brain examination using optimized wave-CAIPI encoding. *J Magn Reson Imaging*. 2019;50:961-974.
- Gagoski BA, Bilgic B, Eichner C, et al. RARE/Turbo spin echo imaging with simultaneous multislice wave-CAIPI. *Magn Reson Med*. 2015;73:929-938.
- Schwarz JM, Pracht ED, Brenner D, Reuter M, Stöcker T. GRAPPA reconstructed wave-CAIPI MP-RAGE at 7 Tesla. *Magn Reson Med*. 2018;80:2427-2438.
- Lustig M, Donoho D, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magn Reson Med*. 2007;58:1182-1195.
- Haldar J. Low-rank modeling of local k-space neighborhoods (LORAKS) for constrained. *IEEE Trans Image Process*. 2013;33:1-27.
- Bilgic B, Ye H, Wald LL, Setsompop K. Optimized CS-Wave imaging with tailored sampling and efficient reconstruction. In Proceedings of the 24th Annual Meeting ISMRM, Singapore. 2016 (p. 612).
- Kim TH, Bilgic B, Polak D, Setsompop K, Haldar JP. Wave-LORAKS: combining wave encoding with structured low-rank

- matrix modeling for more highly accelerated 3D imaging. *Magn Reson Med.* 2019;81:1620-1633.
16. Block KT, Uecker M, Frahm J. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magn Reson Med.* 2007;57:1086-1098.
 17. Knoll F, Bredies K, Pock T, Stollberger R. Second order total generalized variation (TGV) for MRI. *Magn Reson Med.* 2011;65:480-491.
 18. Hyun CM, Kim HP, Lee SM, Lee S, Seo JK. Deep learning for undersampled MRI reconstruction. *Phys. Med. Biol.* 2018;63:135007.
 19. Polak D, Gong E, Cauley SF, et al. Faster 3D brain scans using Wave-CAIPI encoding and Deep Learning denoising. Paper presented at: ISMRM Workshop on Machine Learning; 2018; Pacific Grove, CA.
 20. Bilgic B, Goyal VK, Adalsteinsson E. Multi-contrast reconstruction with Bayesian compressed sensing. *Magn Reson Med.* 2011;66:1601-1615.
 21. Gong E, Huang F, Ying K, Wu W, Wang S, Yuan C. PROMISE: parallel-imaging and compressed-sensing reconstruction of multi-contrast imaging using sharable information. *Magn Reson Med.* 2015;73:523-535.
 22. Knoll F, Holler M, Koesters T, Otazo R, Bredies K, Sodickson DK. Joint MR-PET reconstruction using a multi-channel image regularizer. *IEEE Trans Med Imaging.* 2017;36:1-16.
 23. Chatnuntawech I, Martin A, Bilgic B, Setsompop K, Adalsteinsson E, Schiavi E. Vectorial total generalized variation for accelerated multi-channel multi-contrast MRI. *Magn Reson Imaging.* 2016;34:1161-1170.
 24. Gong E, Zaharchuk G, Pauly J. Improving the PI+CS Reconstruction for Highly Undersampled Multi-contrast MRI using Local Deep Network. Paper presented at: Proceedings of the 25th Annual Meeting of the ISMRM 2017. Honolulu, Hawaii. Abstract 5663.
 25. Dar SU, Yurt M, Karacan L, Erdem A, Erdem E, Cukur T. Image synthesis in multi-contrast MRI with conditional generative adversarial networks. *IEEE Trans Med Imaging.* 2019; 38:2375-2388.
 26. Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proc IEEE Int Conf Comput Vis.* 2017;2017:2242-2251.
 27. Akçakaya M, Moeller S, Weingärtner S, Uğurbil K. Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: database-free deep learning for fast imaging. *Magn Reson Med.* 2019;81:439-453.
 28. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. *Nature.* 2018;555:487.
 29. Hammernik K, Klatzer T, Kobler E, et al. Learning a variational network for reconstruction of accelerated MRI data. *Magn Reson Med.* 2018;79:3055-3071.
 30. Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans. Med. Imaging.* 2018;37:491-503.
 31. Mardani M, Gong E, Cheng JY, et al. Deep generative adversarial neural networks for compressive sensing MRI. *IEEE Trans Med Imaging.* 2019;38:167-179.
 32. Chen F, Cheng JY, Taviani V, et al. Data-driven self-calibration and reconstruction for non-Cartesian wave-encoded single-shot fast spin echo using deep learning. *J. Magn. Reson.* 2019;26871.
 33. Cauley SF, Setsompop K, Bilgic B, Bhat H, Gagoski B, Wald LL. Autocalibrated wave-CAIPI reconstruction: joint optimization of k-space trajectory and parallel imaging reconstruction. *Magn Reson Med.* 2017;78:1093-1099.
 34. Roth S, Black MJ. Fields of experts. *Int J Comput Vis.* 2009;82:205-229.
 35. Bilgic B, Kim TH, Liao CU, et al. Improving parallel imaging by jointly reconstructing multi-contrast data. *Magn Reson Med.* 2018;80:619-632.
 36. Mugler JP, Bao S, Mulkern RV, et al. Optimized single-slab three-dimensional spin-Echo MR imaging of the brain. *Radiology.* 2000;216:891-899.
 37. Jenkinson M, Bannister PR, Brady M, Smith SM. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage.* 2002;17:825-841.
 38. Uecker M, Lai P, Murphy MJ, et al. ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. *Magn Reson Med.* 2014;71:990-1001.
 39. Pock T, Sabach S. Inertial proximal alternating linearized minimization (iPALM) for nonconvex and nonsmooth problems. *SIAM J Imaging Sci.* 2016;9:1756-1787.
 40. Maclaren J, Herbst M, Speck O, Zaitsev M. Prospective motion correction in brain imaging: a review. *Magn Reson Med.* 2013;69:621-636.
 41. Haskell MW, Cauley SF, Wald LL. TArgeted motion estimation and reduction (TAMER): data consistency based motion mitigation for mri using a reduced model joint optimization. *IEEE Trans Med Imaging.* 2018;37:1253-1265.
 42. Haskell MW, Cauley SF, Bilgic B, et al. Network accelerated motion estimation and reduction (NAMER): convolutional neural network guided retrospective motion correction using a separable motion model. *Magn Reson Med.* 2019;82:1452-1461.
 43. Skare S, Sprenger T, Norbeck O, et al. A 1-minute full brain MR exam using a multicontrast EPI sequence. *Magn Reson Med.* 2018;79:3045-3054.
 44. Liao C, Cao X, Cho J, Zhang Z, Setsompop K, Bilgic B. Highly efficient MRI through multi-shot echo planar imaging. *arXiv:1908.00983*, 2019.
 45. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage.* 2016;125:1063-1078.
 46. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process.* 2004;13:600-612.
 47. Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. *Adv Neural Inf Process Syst.* 2014;3:2672-2680.
 48. Hammernik K, Knoll F, Sodickson D, Pock T. L2 or Not L2: Impact of loss function design for deep learning MRI reconstruction. Paper presented at: Proceedings of the International Society of Magnetic Resonance in Medicine; 2017. Honolulu, HI. Abstract 0687.
 49. Hammernik K, Kobler E, Pock T, Recht MP, Sodickson DK, Knoll F. Variational adversarial networks for accelerated MR image reconstruction. Paper presented at: Proceedings of the International Society of Magnetic Resonance in Medicine; 2018; Paris, France. Abstract 1091.
 50. Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for image restoration with neural networks. *IEEE Trans Comput Imaging.* 2016;3:47-57.

SUPPORTING INFORMATION

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

TABLE S1 Acquisition parameters for T1w, T2w and T2-FLAIR-weighted SPACE

TABLE S2 Quantitative metrics (NRMSE, SSIM and PSNR) computed across 36 slices (test subject #2) are provided for T2-FLAIR, T2w and T1w reconstructions at $R = 6$ -fold acceleration

FIGURE S1 Sampling trajectory for BPE at 1 mm isotropic resolution and $R = 6$ -fold acceleration (note, only modulation along k_y is shown, but analogous for k_z)

How to cite this article: Polak D, Cauley S, Bilgic B, et al. Joint multi-contrast variational network reconstruction (jVN) with application to rapid 2D and 3D imaging. *Magn Reson Med*. 2020;84:1456–1469. <https://doi.org/10.1002/mrm.28219>

APPENDIX

Figure A1 compares *SENSE*, *VN*, and *jVNc* to a combined non-linear *PI-CS* reconstruction technique ($R = 6$ -fold acceleration, uniform sampling without integrated ACS). We utilized Total Generalized Variation (TGV)¹⁷ for the single contrast (*SC-TVG*) reconstruction and Vectorial Total Generalized Variation for the multi-contrast (*MC-TGV*)²³ reconstruction across T1w, T2w and T2-FLAIR-weighted data. *SC-TVG* significantly reduced aliasing artifacts and NRMSE when compared to *SENSE*. Moreover, reconstructing all contrasts jointly (*MC-TVG*) slightly improved the image quality and enabled better preservation of fine details. This is best seen in the zoom-in, where the spatial resolution at the intersection between CSF and white matter was better retained. Nevertheless, learnt regularization in *VN* still outperformed both *PI-CS* techniques at $R = 6$ -fold acceleration based upon the quantitative metrics. The overall best image quality was achieved by *jVN*, which enabled 30% NRMSE improvement over *VN*.

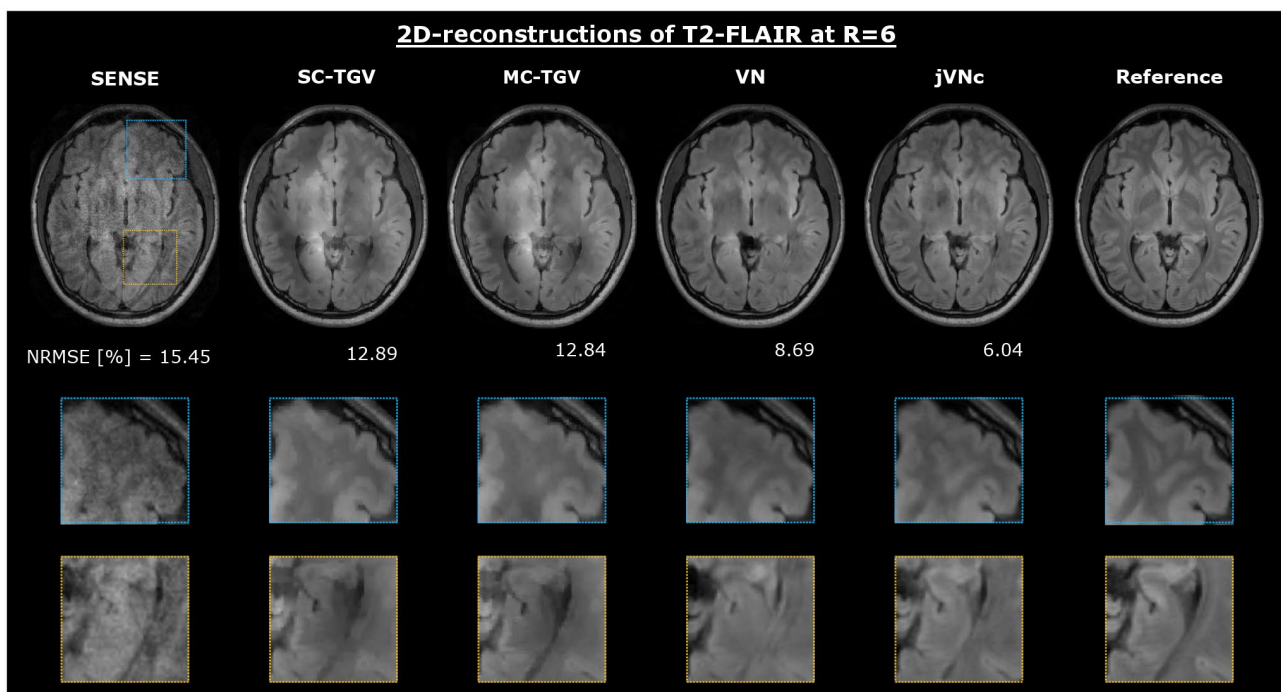


FIGURE A1 Comparison of *VN* and *jVNc* against single and multi-contrast *PI-CS* reconstructions with TGV regularization. NRMSE was computed for a single slice