Death by Retrospective Undersampling - Caveats and Solutions for Learning-Based MRI Reconstructions

Junaid R. Rajput[1,2], Simon Weinmueller[1], Jonathan Endres[1], Peter Dawood[1], Florian Knoll[3], Andreas Maier[2], and Moritz Zaiss[1,3]

[1] Institute of Neuroradiology, University Hospital Erlangen, Erlangen, Germany
[2] Pattern Recognition Lab Friedrich-Alexander-University Erlangen-Nürnberg
junaid.rajput@fau.de
[3] Department of Artificial Intelligence in Biomedical Engineering Friedrich Alexander University of Erlangen–Nuernberg, Erlangen, Germany

Abstract. This study challenges the validity of retrospective undersampling in MRI data science by analysis via an MRI physics simulation. We demonstrate that retrospective undersampling, a method often used to create training data for reconstruction models, can inherently alter MRI signals from their prospective counterparts. This arises from the sequential nature of MRI acquisition, where undersampling post-acquisition effectively alters the MR sequence and the magnetization dynamic in a non-linear fashion. We show that even in common sequences, this effect can make learning-based reconstructions unreliable. Our simulation provides both, (i) a tool for generating accurate prospective undersampled datasets for analysis of such effects, or for MRI training data augmentation, and (ii) a differentiable reconstruction operator that models undersampling correctly. The provided insights are crucial for the development and evaluation of AI-driven acceleration of diagnostic MRI tools.

Keywords: Retrospective undersampling · Magnetization · Bloch simulation.

1 Introduction

Acceleration in MRI technology has progressed through three pivotal advancements: (i) the introduction of rapid sequence methodologies like fast low angle shot (FLASH) [6] and turbo spin echo (TSE) [2], (ii) the adoption of parallel imaging techniques enabling partial k-space acquisitions [4], and (iii) the development of compressed sensing strategies, which permit further k-space undersampling through iterative reconstruction, leveraging sparse domains and ensuring data consistency [9]. The latter advancement is notably compatible with machine learning, facilitating the fourth transition to learning-based reconstruction methods that significantly enhance the speed of clinical MRI.

Learning-based reconstruction models, such as U-NETs [15] or variational networks (VN) [7], necessitate specially prepared training data [14]. This is typi-

cally generated from fully sampled k-space datasets, from which the target imaging data is derived. The process involves retrospectively undersampling of fully sampled datasets at various acceleration factors, presupposing that the undersampling does not influence the actual MRI acquisition process.

However, this assumption holds true primarily for sequences operating in a steady-state regime Fig. 1(a,d). Many MRI sequences inherently incorporate relaxation or preparation phases, interleaved with acquisition periods. During these phases, the system remains in a transient state, with magnetization and signal intensity fluctuating not only throughout the acquisition but also as a direct consequence of it Fig. 1(b,c,e,f). Thus, undersampling in such transient or pseudo-steady-state sequences as TSE and FLAIR shortens the acquisition phase and thereby changes the dynamic signal evolution. Consequently, data acquired prospectively with undersampling differ from those acquired retrospectively with undersampling.

This critical aspect is often overlooked or disregarded in the development of learning-based reconstruction algorithms, leading to suboptimal model performance and deviations growing with the acceleration factors.

Within this article, we aim to: (i) illustrate this discrepancy through precise Bloch-Simulation visualization, (ii) demonstrate its impact on the efficacy of learning-based reconstruction models, and (iii) introduce a comprehensive and differentiable Bloch model as an adaptable solution to the challenges posed by undersampling.

2 Methods

2.1 Simulation Setup

For our study, we use the differentiable MRI simulation from the MR-zero framework [8]. It utilizes the improved implementation of the extended phase graph algorithm [16] called phase distribution graph (PDG), which has been recently validated as a realistic and comprehensive MRI simulation by analytical solutions, other simulation frameworks and in vivo measurements [5]. It includes description of T_1, T_2, T_2^*, PD and diffusion processes as well as B_0 and B_1 field inhomogeneities [5, 10]. The MR-zero simulation is implemented in pytorch and generates the full MR signal proportional to the transverse magnetization M_{xy} in every voxel without any simplifications and including the full sequence history t(N) for any (undersampled) sequence (seq), the magnetization in each voxel changes throughout the simulation and for any sampling point n in measurement m.

$$M_{xy}(N = m * n) = Bloch(seq, t(N)) \tag{1}$$

Due to auto-differentiation in pytorch, this operator can be used for model based reconstruction [12]. The assumption of invariance against undersampling makes the conceptual error that the signal acquisition has no effect on the current or subsequent acquisitions. This can be interpreted as assuming the signal follows an equation which is independent of any previous acquisition Fig. 1(c).

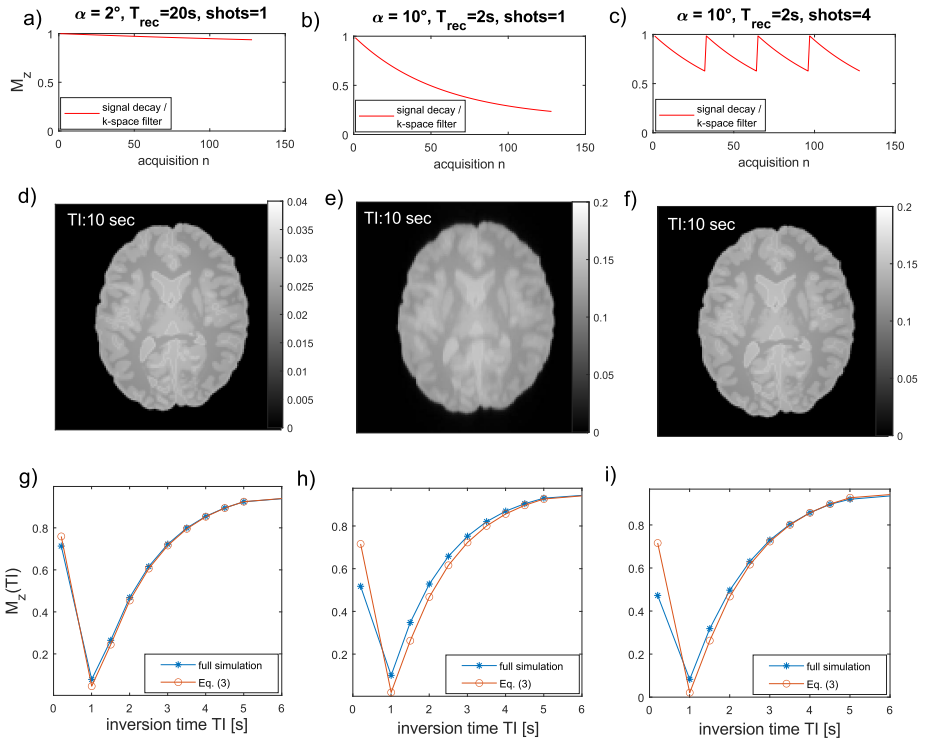

Fig. 1. (a-c) IR-FLASH sequence readout phase for different flip angle and shots, leading to different blurriness in the prepared images (d-f). In addition, the readout and limited recovery time alter the dynamic of the MR signal (g-i), which deviates from the ideal equation (3).

To challenge this assumption we choose a common sequence that has such a signal equation, namely an inversion recovery magnetization prepared rapid acquisition of gradient echoes (MPRAGE). In 2D, which is sufficient for our argumentation this sequence is also known as an inversion recovery prepared fast low angle shot (IR-FLASH) sequence [3].

Assuming a pseudo-steady state, where each readout reaches a steady state at every time point, the equation, independent of N, can be stated as follows:

$$M_z(TI_m) = M_0 \left(1 - \left(2 - \frac{1 - e^{-T_R/T_1}}{1 - \cos(\alpha) \cdot e^{-T_R/T_1}} \cdot e^{-T_{REC}/T_1}\right) \cdot e^{-TI_m/T_1}\right) \quad (2)$$

This equation describes the longitudinal magnetization $M_z(TI)$ at a time TI after the inversion. Here, M_0 is the equilibrium magnetization, TI_m is the mth inversion time after the inversion pulse and before the FLASH readout which acquires N acquisitions. T_R is the repetition time of n acquisitions within each FLASH sequence, T_1 is the longitudinal relaxation time, α is the flip angle used in

the FLASH sequence, and T_{REC} is the recovery time after the FLASH sequence before the next prepared scan m.

The signal provided by the FLASH readout is then given by:

$$M_{xy} = M_z(TI, T_{REC}) \sin(\alpha) \tag{3}$$

By comparing eq. 3 with the simulation (eq. 1), we can show when the assumption of undersampling invariance of the signal is violated. The IR-FLASH used has a flip angle of 10°, the recovery time $T_{REC} = 2$ seconds and 11 inversion times $TI = [10.0, 5.0, 4.5, 4.0, 3.5, 3.0, 2.5, 2.0, 1.5, 1.0, 0.2]$. $TI = 100$ was assumed to be fully relaxed and used as a normalizing scan.

2.2 Unrolled Variational Network

To test the effect of prospective undersampling on neural networks based on retrospectively undersampled data, we used the state of the art unrolled variational network (VN) [7]. The unrolled VN used in the study uses 10 iterative blocks. Specifically, the regularization module in each iterative block contains 36 sets of 3D kernels (size = 7 x 7 x 7) and corresponding activation functions, each of which is characterized by a weighted combination of 31 Gaussian radial basis functions [7].

2.3 Training and Evaluation

The training data consisted of synthetic brain samples based on the BrainWeb database [1]. The fuzzy model segments were filled with in vivo-like tissue parameters. It was assumed that B_0 and B_1 had no inhomogeneities. A total of 15 subject volumes, each with 70 slices, were used for training. In addition, 3 subject volumes, also with 70 slices each, were determined for validation in order to fine-tune the model parameters. Finally, 2 subject volumes, were used as a test dataset. The VN was trained with the mean square error (MSE) for the real and imaginary part of the 5th fold undersampled image, using the target as a fully sampled complex image. The evaluation was based on structural similarity (SSIM) and root mean square error (RMSE).

Table 1. Performance Metrics for VN trained on retrospective undersampeld data.

Test Configuration	SSIM	RMSE	MAE
Retrospectively undersampled	0.98987	0.00347	0.00201
Prospectively undersampled	0.94185	0.01437	0.00767

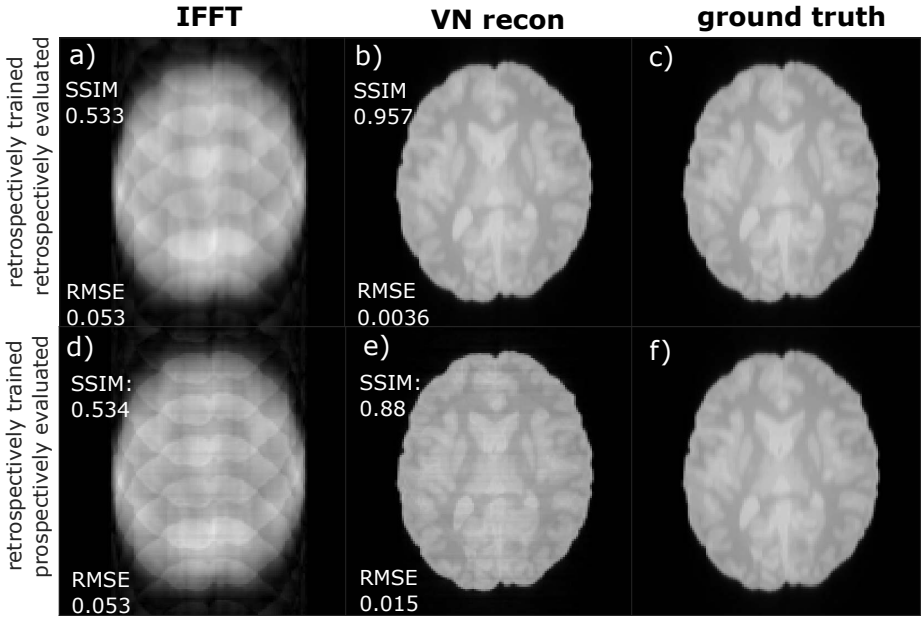

Fig. 2. Variational network reconstructions after training on retrospectively undersampled data. (a,b) Magnitude image of the retrospective undersampled input and the corresponding VN reconstruction. (c,d) Magnitude image of the prospective undersampled input and the corresponding VN reconstruction. SSIM and RMSE were calculated with respect to the fully sampled ground truth magnitude image.

Table 2. Statistical validation of differences in model performance: comparison of SSIM, RMSE and MAE between tests on retrospective and prospective undersampled data.

Metric	**T-statistic**	**p-value**
SSIM	90.1385	4.67×10^{-139}
RMSE	-92.6020	6.78×10^{-141}
MAE	-52.8006	3.80×10^{-103}

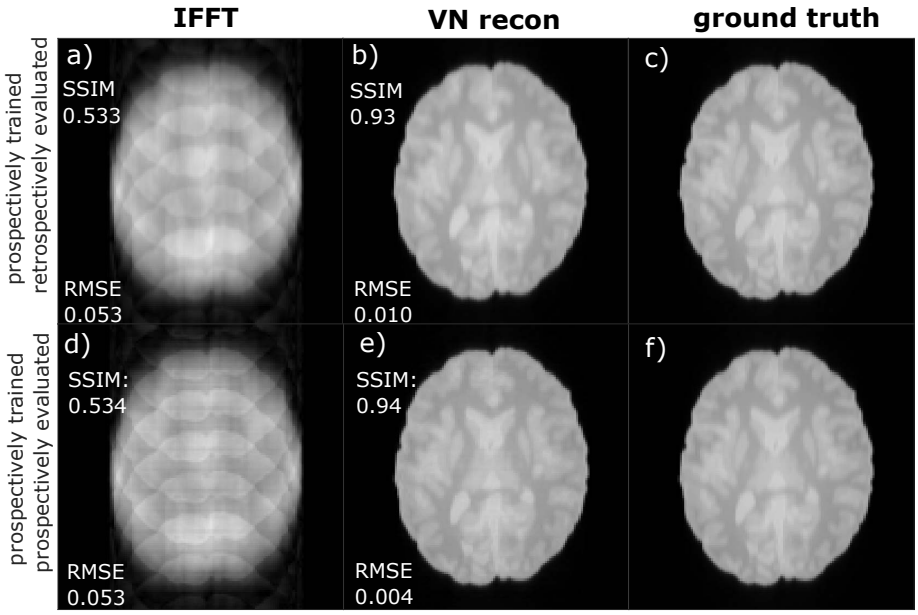

Fig. 3. Variational network reconstructions after training on prospectively undersampled data. (a,b) Magnitude image of the retrospective undersampled input and the corresponding VN reconstruction. (c,d) Magnitude image of the prospective undersampled input and the corresponding VN reconstruction. SSIM and RMSE were calculated with respect to the fully sampled ground truth magnitude image.

3 Results

Fig. 1(g-i) shows the comparison of the signal equation (eq. 3) and the MR simulation (eq. 1) for three different cases. (i) For long T_{REC}, and low flip angle, where both match and are independent of the acquisition. (ii) For short T_{REC} and high flip angle, where they deviate, and the simulated signal changes with acquisition. (iii) like (ii) but for four FLASH shots, which reflects the same change in dynamic as undersampling by factor 4. This clearly has both short term blurring effects on the current acquisition Fig. 1(e,f), but also long term history effects on later preparations and acquisitions.

Fig. 2 shows the impact of neglecting the effect of retrospective undersampling by training the VN on retrospectively undersampled data and inferring it for retrospective and prospective 5-fold undersampled data. The performance for retrospectively undersampled data (Fig. 2a) is close to ground truth fully sampled image Fig. 2(b,c), while the reconstruction for prospectively undersampled (Fig. 2d) examples shows visible artifacts and poor performance (Fig. 2e), as the RMSE increases by a factor of 5 and the SSIM decreases by 8%. Evaluation metric were computed for the magnitude input image (IFFT) and reconstructed VN output with respect to the fully sampled ground truth. Table 1 summarizes

the comparison of performance metrics (SSIM, RMSE and MAE) for a model trained on retrospective data and tested on both retrospective and prospective data for complete test set. It clearly demonstrates higher SSIM and lower RMSE and MAE for retrospective undersampled data. The hypothesis tests reinforce these observations, providing strong statistical backing to the observed differences in performance (Table 2). Fig. 3 shows that this can be resolved when training is done with the right dynamics (multi shot data or simulated prospective undersampled data).

Fig. 4(d-f) shows the deviation of the full T_1 quantification of the multi-IR FLASH sequence via MR-zero for the retrospectively undersampled target, as the fitted data were prospectively undersampled. In contrast, quantification for the prospective setting, i.e. the target and data were both prospectively undersampled, the iterative fit yields an RMSE of only 0.013 (Fig. 4(a-c)).

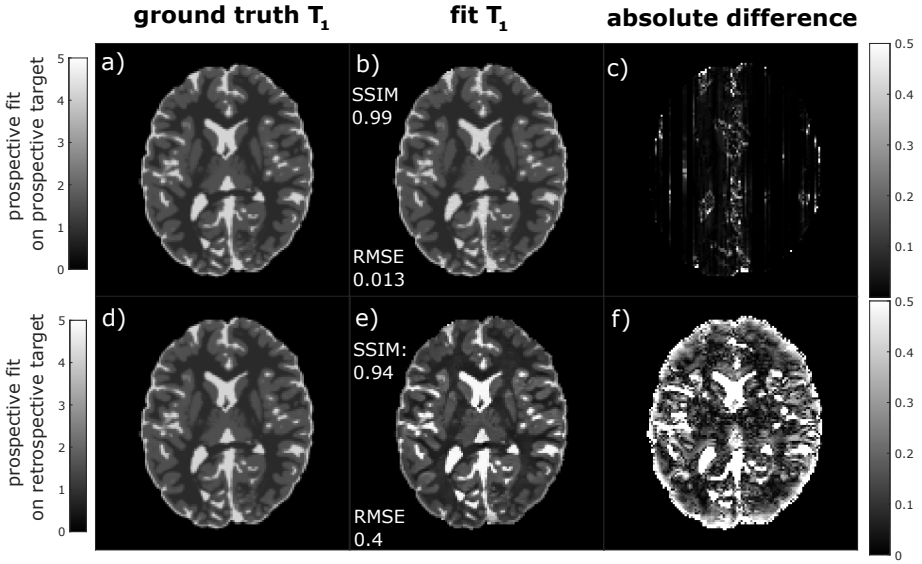

Fig. 4. T_1 quantification via MR-zero framework. (a-c) Prospective setting, fitted data and target were both prospectively undersampled. (d-f) Retrospective setting, fitted data was prospectively undersampled but target was retrospectively undersampled.

4 Discussion

In this article we showed that the common practice of retrospective undersampling of existing data [15, 17] to generate training data for learning-based MRI reconstruction models have to be done with caveats. The performance of models trained via retrospective undersampled data when applied to prospective data

remains therefore unclear, and, as the deviation from prospective data increases with the undersampling factor, highly accelerated methods might currently not yet be at their theoretical limit. Our simulation approach allows to identify such mismatches, correct data simulation, and even reconstruction with the complete and differentiable MRI operator. Also, without simulation, matching training data can be acquired in vivo by acquiring prospective undersampled data together with a fully sampled data, or in the correct segmentation mode so that the dynamics are identical, such as in [11]. Still such scans can be very time consuming especially if several undersampling factors need to be tested and acquired. Thus, data generation or augmentation by our simulation model depict a highly efficient solution. While Bloch-model-based reconstructions exist [13], the insights of Bloch dynamics rarely applied in learning-based approaches. Our simulation lowers here the threshold as it is implemented within pytorch, a native machine learning framework.

5 Conclusion

Learning-based MRI reconstruction models may die a "death by retrospective undersampling" if retrospective undersampling would compromise the dynamics of prospective acquisition, which is the case for common clinical MRI sequences. This might limit current highly accelerated methods in their performance. The differentiable MR-zero simulation provides a solution as correct reconstruction operator or generator of correctly simulated undersampled training data.

Acknowledgments. This work was supported by Deutsche Forschungsgemeinschaft, Grant/Award Number: ZA8147-1.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. BrainWeb: Simulated Brain Database, https://brainweb.bic.mni.mcgill.ca/brainweb/
2. Dang, H.N., Endres, J., Weinmüller, S., Glang, F., Loktyushin, A., Scheffler, K., Doerfler, A., Schmidt, M.A., Maier, A.K., Zaiss, M.: Mr-zero meets rare mri: Joint optimization of refocusing flip angles and neural networks to minimize t2-induced blurring in spin echo sequences. Magnetic Resonance in Medicine **90**, 1345 – 1362 (2023), https://api.semanticscholar.org/CorpusID:259250074
3. Deichmann, R., Haase, A.: Quantification of t1 values by snapshot-flash nmr imaging. Journal of Magnetic Resonance **96**, 608–612 (1992), https://api.semanticscholar.org/CorpusID:55465696
4. Deshmane, A., Gulani, V., Griswold, M.A., Seiberlich, N.: Parallel mr imaging. Journal of Magnetic Resonance Imaging **36**(1), 55–72 (Jun 2012). https://doi.org/10.1002/jmri.23639, http://dx.doi.org/10.1002/jmri.23639

5. Endres, J., Weinmüller, S., Dang, H.N., Zaiss, M.: Phase distribution graphs for fast, differentiable, and spatially encoded Bloch simulations of arbitrary MRI sequences. Magnetic Resonance in Medicine (2023). https://doi.org/10.1002/mrm.30055
6. Haase, A., Frahm, J., Matthaei, D., Hänicke, W., Merboldt, K.D.: Flash imaging: rapid nmr imaging using low flip-angle pulses. 1986. Journal of magnetic resonance **213 2**, 533–41 (1986), https://api.semanticscholar.org/CorpusID:36983792
7. Hammernik, K., Klatzer, T., Kobler, E., Recht, M.P., Sodickson, D.K., Pock, T., Knoll, F.: Learning a variational network for reconstruction of accelerated mri data. Magnetic Resonance in Medicine **79** (2017), https://api.semanticscholar.org/CorpusID:3815411
8. Loktyushin, A., Herz, K., Dang, N.P., Glang, F., Deshmane, A., Weinmüller, F., Doerfler, A., Scholkopf, B., Scheffler, K., Zaiss, M.: Mrzero: Fully automated invention of mri sequences using supervised learning (2020), https://api.semanticscholar.org/CorpusID:211076178
9. Lustig, M., Donoho, D.L., Santos, J.M., Pauly, J.M.: Compressed sensing mri. IEEE Signal Processing Magazine **25**, 72–82 (2008), https://api.semanticscholar.org/CorpusID:945906
10. Malik, S.J., Teixeira, R.P.A.G., Hajnal, J.V.: Extended phase graph formalism for systems with magnetization transfer and chemical exchange. arXiv: Medical Physics (2017), https://api.semanticscholar.org/CorpusID:119443320
11. McClymont, D., Teh, I., Whittington, H.J., Grau, V., Schneider, J.E.: Prospective acceleration of diffusion tensor imaging with compressed sensing using adaptive dictionaries. Magnetic Resonance in Medicine **76**(1), 248–258 (Jul 2016). https://doi.org/10.1002/mrm.25876, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4869836/
12. Rajput, J.R., Möhle, T.A., Fabian, M.S., Mennecke, A., Sembill, J.A., Kuramatsu, J.B., Schmidt, M., Dörfler, A., Maier, A., Zaiss, M.: Physics-informed conditional autoencoder approach for robust metabolic cest mri at 7t. In: Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., Taylor, R. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. pp. 449–458. Springer Nature Switzerland, Cham (2023)
13. Scholand, N., Wang, X., Roeloffs, V., Rosenzweig, S., Uecker, M.: Quantitative mri by nonlinear inversion of the bloch equations. Magnetic Resonance in Medicine **90**(2), 520–538 (2023). https://doi.org/https://doi.org/10.1002/mrm.29664, https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.29664
14. Shimron, E., Tamir, J.I., Wang, K., Lustig, M.: Implicit data crimes: Machine learning bias arising from misuse of public data. Proceedings of the National Academy of Sciences of the United States of America **119** (2022), https://api.semanticscholar.org/CorpusID:247597655
15. Singh, D., Monga, A., de Moura, H.L., Zhang, X., Zibetti, M.V.W., Regatte, R.R.: Emerging Trends in Fast MRI Using Deep-Learning Reconstruction on Undersampled k-Space Data: A Systematic Review. Bioengineering **10**(9), 1012 (Sep 2023). https://doi.org/10.3390/bioengineering10091012, https://www.mdpi.com/2306-5354/10/9/1012, number: 9 Publisher: Multidisciplinary Digital Publishing Institute
16. Weigel, M.: Extended phase graphs: Dephasing, rf pulses, and echoes - pure and simple. Journal of Magnetic Resonance Imaging **41** (2015), https://api.semanticscholar.org/CorpusID:6129394

17. Yiasemis, G., Sánchez, C.I., Sonke, J.J., Teuwen, J.: On retrospective k-space subsampling schemes for deep MRI reconstruction. Magnetic Resonance Imaging **107**, 33–46 (Apr 2024). https://doi.org/10.1016/j.mri.2023.12.012, https://www.sciencedirect.com/science/article/pii/S0730725X23002199