Supporting Information: Dynamic MRI using model-based deep learning and SToRM priors: MoDL-SToRM

The supplementary material consists of more elaborate comparisons of the proposed framework. We also compare the proposed scheme with MoDL in S.3-S.5. The MoDL framework was originally introduced for static images. For clarity, we now briefly review MoDL in the context of free breathing cardiac image reconstruction.

Brief review of MoDL (1): . We note that the MoDL algorithm motivated the development of the MoDL-SToRM framework, proposed in the paper. MoDL formulates the recovery of the images from the k-space data as

$$\mathcal{C}(\mathbf{X}) \quad = \quad \underbrace{\|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_2^2}_{\text{data consistency}} \quad + \quad \frac{\lambda_1}{2}\underbrace{\|\mathcal{N}_{\mathbf{w}}(\mathbf{X})\|^2}_{\text{CNN prior}} \qquad [1]$$

Here, $\mathcal{A}$ is the multi-channel Fourier sampling operator, which includes coil sensitivity weighting. $\mathcal{N}_{\mathrm{w}}$ is a 3-D CNN based estimator that estimates the noise and alias patterns in the dataset from local neighborhoods of the 2D+time dataset; $\|\mathcal{N}_{\mathbf{w}}(\mathbf{x})\|^2$ is a measure of the alias/noise contribution in the dataset $\mathbf{X}$. The *denoised signal* can thus be estimated from the data $\mathbf{X}$ as $\mathcal{D}_{\mathbf{w}}(\mathbf{X}) = (\mathcal{I}-\mathcal{N}_{\mathrm{w}})(\mathbf{X}) = \mathbf{X}-\mathcal{N}_{\mathbf{w}}(\mathbf{X})$.

We consider temporary variable $\mathbf{Y} = \mathcal{D}_{\mathbf{w}}(\mathbf{X})$ and rewrite [1] as: $\mathcal{C}(\mathbf{X}) = \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_2^2 + \frac{\lambda_1}{2}\|\underbrace{\mathbf{X}-\mathbf{Y}}_{\mathcal{N}_w(\mathbf{X})}\|^2$
Minimizing the objective with respect to $\mathbf{X}$, assuming variables $\mathbf{Y}$ to be fixed and determined from the previous iterations yields the following alternating optimization: (see Fig. S1)

$$\mathbf{Y}_n \quad = \quad \mathcal{D}_{\mathbf{w}}(\mathbf{X}_n) \qquad [2]$$

$$\mathbf{R}_n \quad = \quad (\mathcal{A}^*(\mathbf{B}) + \lambda_1\mathbf{Y}_n) \qquad [3]$$

$$\mathbf{X}_{n+1} \quad = \quad (\mathcal{A}^*\mathcal{A} + \lambda_1\mathbf{I})^{-1}\mathbf{R}_n. \qquad [4]$$

Once the number of iterations is fixed, the network can be unrolled to yield a deep network as in Fig. S.1.(a). The parameters of $\mathcal{D}_{\mathbf{w}}$ and λ_1 are trainable and shared throughout the iterations. We consider reconstructions from 200 frames, corresponding to 8.4sec of acquisition. The total number of trainable parameters were same for MoDL and MoDL-SToRM (with exception of λ_2) strategies.

Comparisons: The elaboration of the Figures 2-5 in the main manuscripts are given in Figures S. 3-5. Likewise, Table S.1 provides includes the comparison with MoDL alone, compared to Table 1 in the main

manuscript.

The poor reconstructions offered by MoDL in figures S.3-5 and Table S1 is expected since MoDL only exploits local redundancies. We note the poor performance of MoDL alone algorithm cannot be attributed to overfitting, as the training and validation loss curves decay monotonically as shown in Fig S2. Note that each image is only sampled with 10 radial lines, which translates to 50 fold undersampling. By contrast, MoDL-SToRM simultaneously exploits local and global redundancies to yield improved results. SToRM facilitates the combination of information from different image frames. With low number of frames, the SToRM alone regularization results in increased errors, which the added MoDL regularization reduces significantly.

LEGENDS OF FIGURES & TABLES IN SUPPORTING INFORMATION

Fig S1: (a) Illustration of MoDL alone implementation used for comparisons in S.3-S.5. The differences between this scheme and current model-based deep-learned schemes are the sharing of the weights across iterations as well as the use of CG blocks to enforce the data-consistency, when complex forward models such as multi-channel sampling is used. (b) i-th iteration of the MoDL: the iterative algorithm alternates between the denoising of the dataset using local CNN block denoted by $\mathcal{D}_w$ and the **DC** block involving conjugate gradients to enforce data-consistency at each iteration. (c) $\mathcal{I} - \mathcal{N}_w = \mathcal{D}_w$, the denoising operator (d) $\mathcal{N}_w$, the noise extractor operator.

Fig S2: Training and validation loss curves for (a) MoDL-alone training (b) proposed: MoDL-SToRM training. The monotonic decay of training and validation error with training iterations show that the models are not over fitted. The sharing of the network parameters across iterations in the MoDL and MoDL-SToRM schemes enable us to keep the number of free parameters low, thus minimizing the risk of overfitting.

Fig S3: Elaboration of Figure 2 in main manuscript. Simulated Dataset: (a) Full view of a single frame from the SToRM reconstruction using 500 frames. Only (red) cropped Myocardium region is shown. (b) Top row: SToRM reconstruction using 500 frames. Following eight rows are four sets of competing reconstructions and corresponding error (w.r.t to top row) images : i) SToRM reconstruction with 100 frames, ii) MoDL with 100 frames, iii) Tikhonov-SToRM reconstruction with 100 frames and iv) proposed with 100 frames. First column is the time profile along a vertical cut across the Myocardium shown in green in (a). Following six columns show three cardiac states at two different respiratory stages. The positions of those two respiratory stages are marked blue and green on the time profiles, in the first column. Three cardiac states are neighboring frames near those two marked time points. The SER (dB) reported in the figure corresponds to the myocardium area.

Fig S4: Elaboration of Figure 3 in main manuscript. Dataset 1: (a) Full view of a single frame from the SToRM reconstruction using 1000 frames. Only (red) cropped Myocardium region is shown. (b) Top row: SToRM reconstruction using 1000 frames. Following eight rows are four sets of competing reconstructions and corresponding error (w.r.t to top row) images : i) SToRM reconstruction with 200 frames, ii) MoDL with 200 frames, iii) Tikhonov-SToRM reconstruction with 200 frames and iv) proposed with 200 frames. First column is the time profile along a vertical cut across the Myocardium shown in green in (a). Following six columns show three cardiac states at two different respiratory stages. The positions of those two respiratory stages are marked blue and green on the time profiles, in the first column. Three cardiac states are neighboring frames near those two marked time points. The SER (dB) reported in the figure corresponds to the myocardium area.

Fig S5: Elaboration of Figure 4 in main manuscript. Dataset 2: (a) Full view of a single frame from the SToRM reconstruction using 1000 frames. Only (red) cropped Myocardium region is shown. (b) Top row: SToRM reconstruction using 1000 frames. Following eight rows are four sets of competing reconstructions and corresponding error (w.r.t to top row) images : i) SToRM reconstruction with 200 frames, ii) MoDL with 200 frames, iii) Tikhonov-SToRM reconstruction with 200 frames and iv) proposed with 200 frames. First column is the time profile along a vertical cut across the Myocardium shown in green in (a). Following six columns show three cardiac states at two different respiratory stages. The positions of those two respiratory stages are marked blue and green on the time profiles, in the first column. Three cardiac states are neighboring frames near those two marked time points. The SER (dB) reported in the figure corresponds to the myocardium area.

Table S1: Elaboration of Table 1 in main manuscript. We compare the reconstruction methods for all test subjects across different recovery metrics. These metrics are reported for the entire field of view. Whereas, the SER (dB) metric in the previous three figures are reported for the myocardium area.

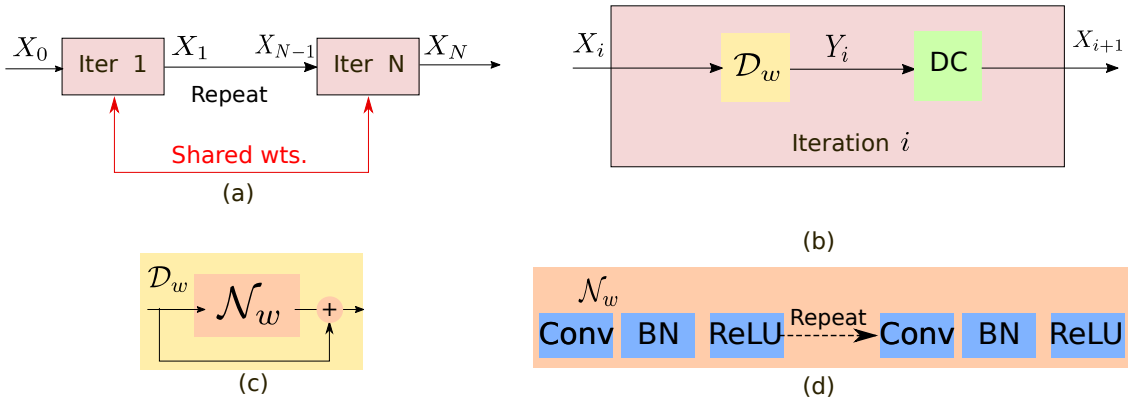


Figure S1: (a) Illustration of MoDL alone implementation used for comparisons in S.3-S.5. The differences between this scheme and current model-based deep-learned schemes are the sharing of the weights across iterations as well as the use of CG blocks to enforce the data-consistency, when complex forward models such as multi-channel sampling is used. (b) i-th iteration of the MoDL: the iterative algorithm alternates between the denoising of the dataset using local CNN block denoted by $\mathcal{D}_w$ and the **DC** block involving conjugate gradients to enforce data-consistency at each iteration. (c) $\mathcal{I} - \mathcal{N}_w = \mathcal{D}_w$, the denoising operator (d) $\mathcal{N}_w$, the noise extractor operator.

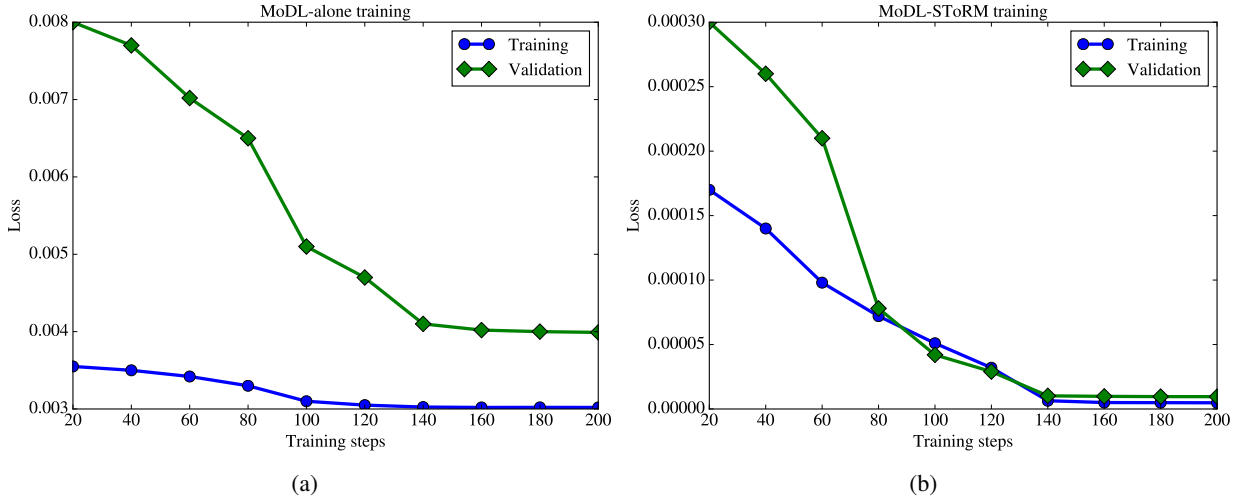
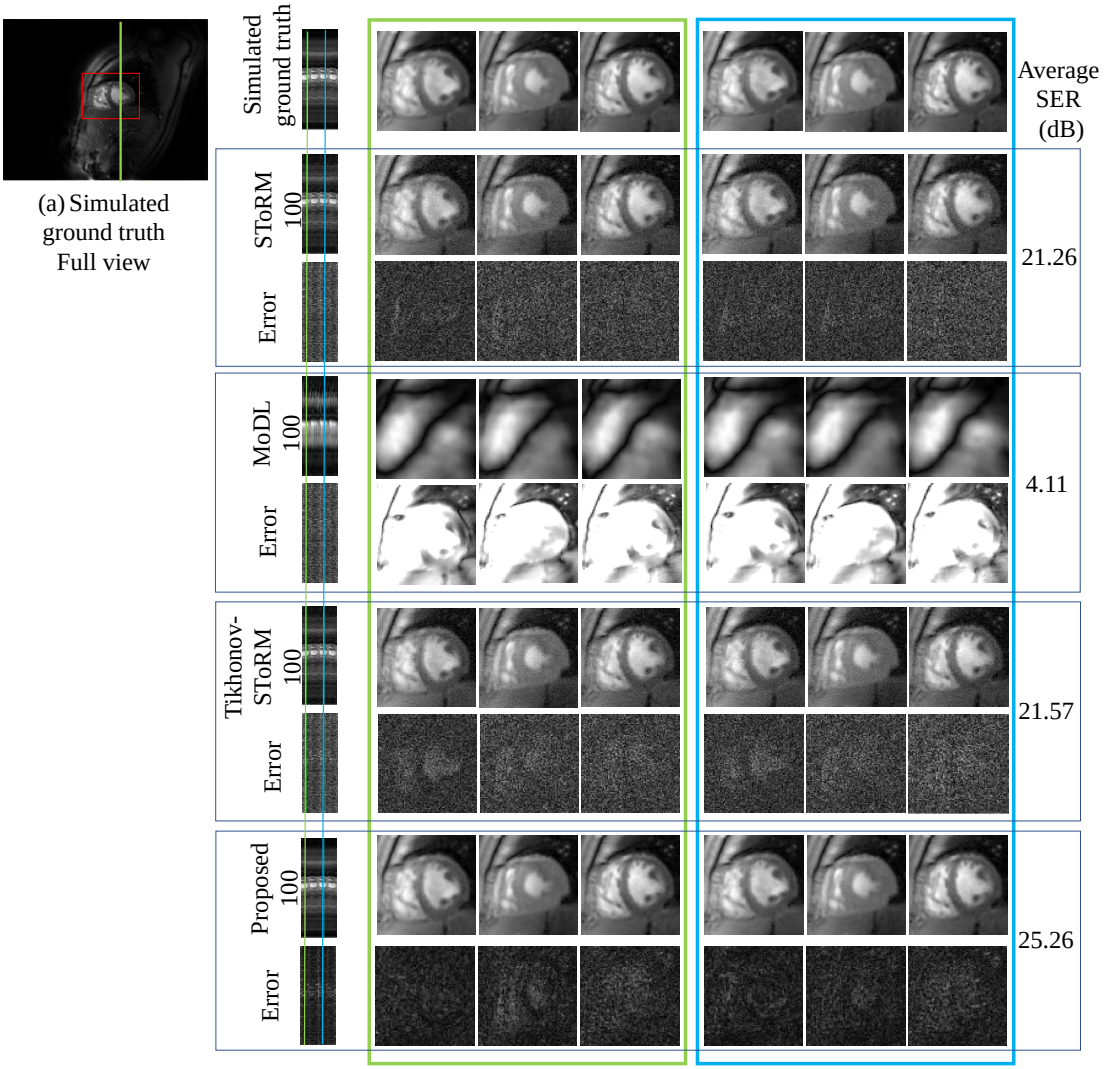


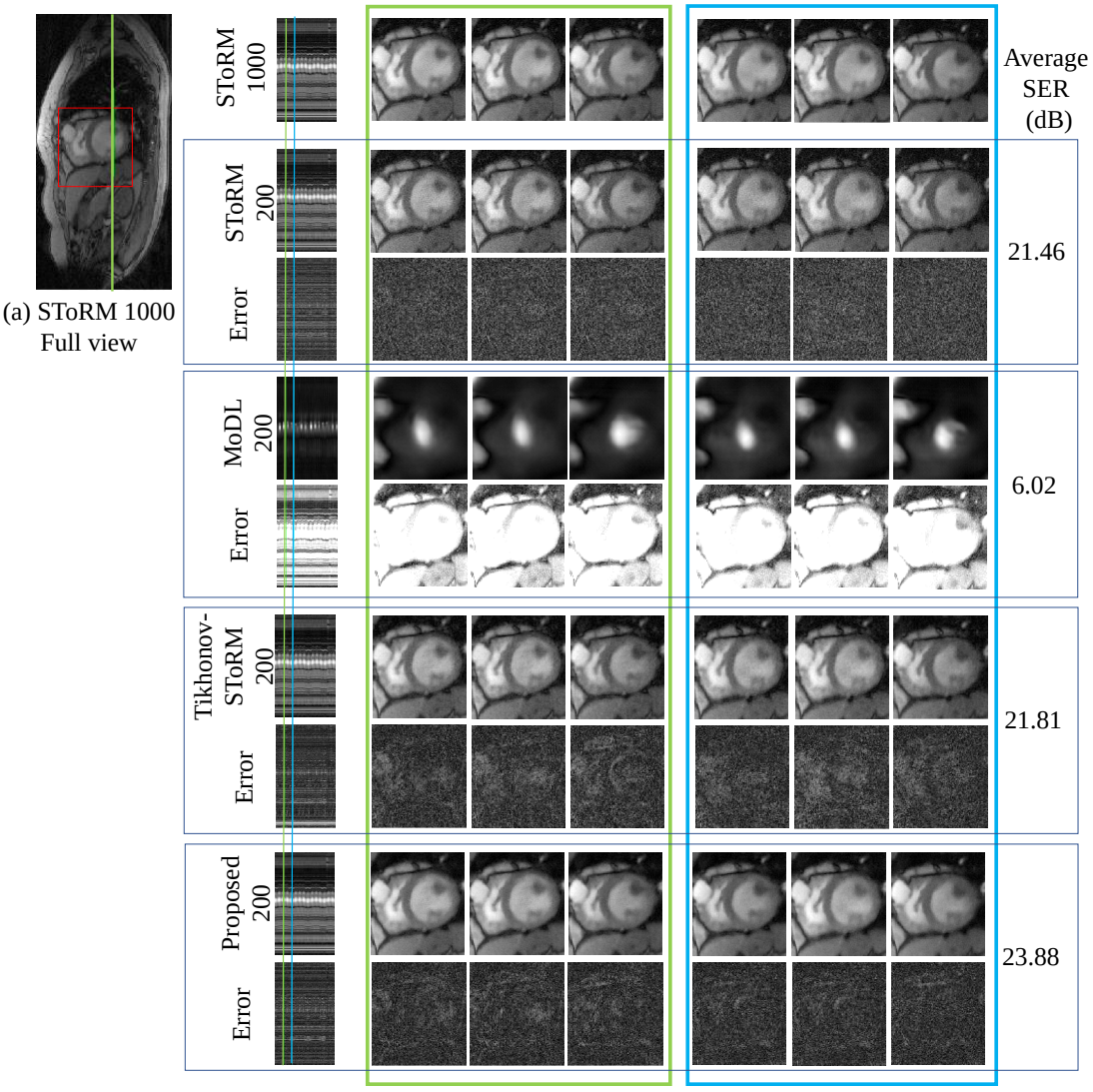
Figure S2: Training and validation loss curves for (a) MoDL-alone training (b) proposed: MoDL-SToRM training. The monotonic decay of training and validation error with training iterations show that the models are not over fitted. The sharing of the network parameters across iterations in the MoDL and MoDL-SToRM schemes enable us to keep the number of free parameters low, thus minimizing the risk of overfitting.



(a) Simulated ground truth Full view

(b) Time profiles, reconstructions and error images at different cardiac & respiratory phases

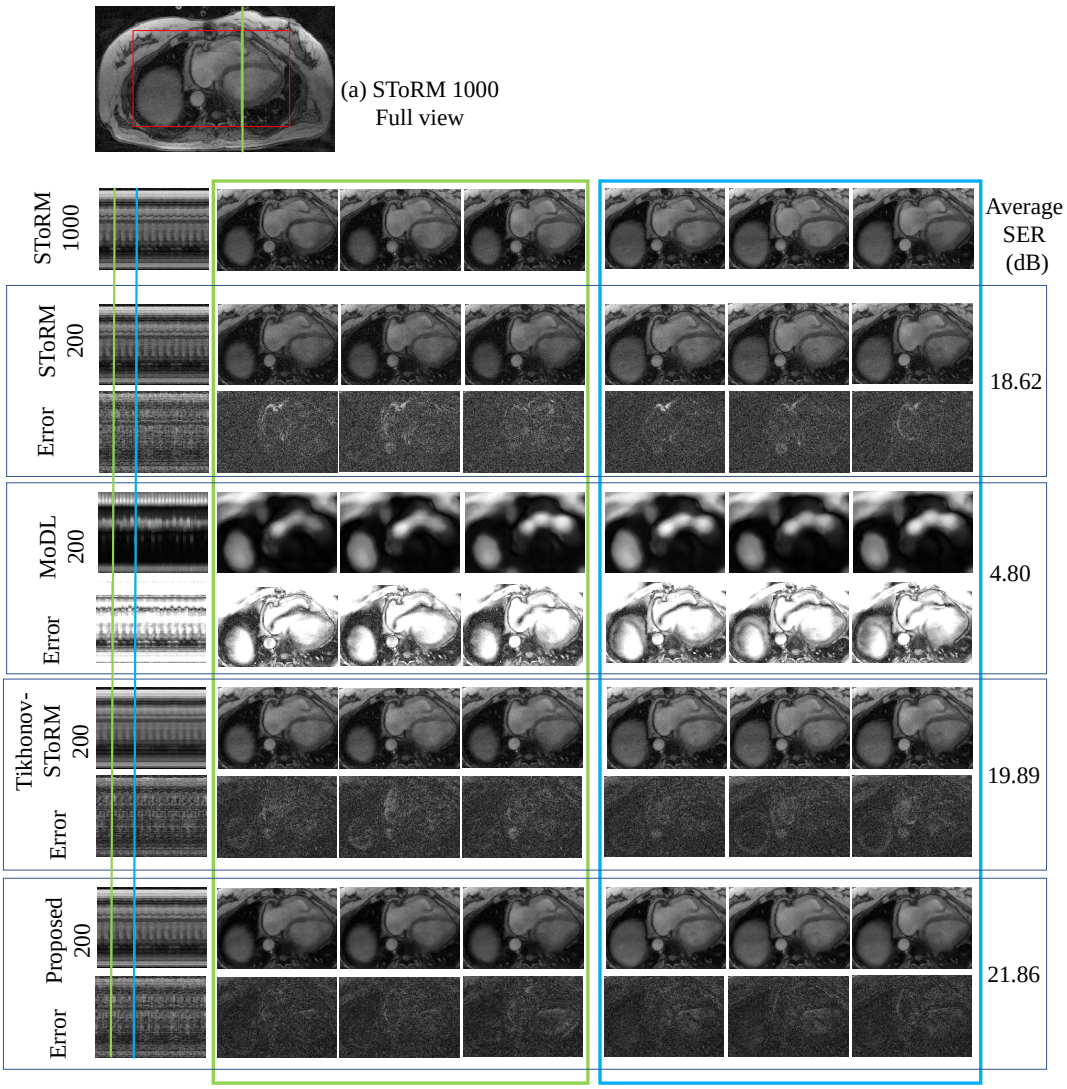
Figure S3: Elaboration of Figure 2 in main manuscript. Simulated Dataset: (a) Full view of a single frame from the SToRM reconstruction using 500 frames. Only (red) cropped Myocardium region is shown. (b) Top row: SToRM reconstruction using 500 frames. Following eight rows are four sets of competing reconstructions and corresponding error (w.r.t to top row) images : i) SToRM reconstruction with 100 frames, ii) MoDL with 100 frames, iii) Tikhonov-SToRM reconstruction with 100 frames and iv) proposed with 100 frames. First column is the time profile along a vertical cut across the Myocardium shown in green in (a). Following six columns show three cardiac states at two different respiratory stages. The positions of those two respiratory stages are marked blue and green on the time profiles, in the first column. Three cardiac states are neighboring frames near those two marked time points. The SER (dB) reported in the figure corresponds to the myocardium area.



(a) STORM 1000 Full view

(b) Time profiles, reconstructions and error images at different cardiac & respiratory phases

Figure S4: Elaboration of Figure 3 in main manuscript. Dataset 1: (a) Full view of a single frame from the SToRM reconstruction using 1000 frames. Only (red) cropped Myocardium region is shown. (b) Top row: SToRM reconstruction using 1000 frames. Following eight rows are four sets of competing reconstructions and corresponding error (w.r.t to top row) images : i) SToRM reconstruction with 200 frames, ii) MoDL with 200 frames, iii) Tikhonov-SToRM reconstruction with 200 frames and iv) proposed with 200 frames. First column is the time profile along a vertical cut across the Myocardium shown in green in (a). Following six columns show three cardiac states at two different respiratory stages. The positions of those two respiratory stages are marked blue and green on the time profiles, in the first column. Three cardiac states are neighboring frames near those two marked time points. The SER (dB) reported in the figure corresponds to the myocardium area.



(b) Time profiles, reconstructions and error images at different cardiac & respiratory phases

Figure S5: Elaboration of Figure 4 in main manuscript. Dataset 2: (a) Full view of a single frame from the SToRM reconstruction using 1000 frames. Only (red) cropped Myocardium region is shown. (b) Top row: SToRM reconstruction using 1000 frames. Following eight rows are four sets of competing reconstructions and corresponding error (w.r.t to top row) images : i) SToRM reconstruction with 200 frames, ii) MoDL with 200 frames, iii) Tikhonov-SToRM reconstruction with 200 frames and iv) proposed with 200 frames. First column is the time profile along a vertical cut across the Myocardium shown in green in (a). Following six columns show three cardiac states at two different respiratory stages. The positions of those two respiratory stages are marked blue and green on the time profiles, in the first column. Three cardiac states are neighboring frames near those two marked time points. The SER (dB) reported in the figure corresponds to the myocardium area.

Dataset	Method	SER (dB)	PSNR (dB)	SSIM
Subject 1	SToRM	16.31	37.31	0.8868
	MoDL	4.06	23.93	0.4381
	Tikhonov-SToRM	18.78	39.77	0.9021
	Proposed	**20.36**	**41.36**	**0.9386**
Subject 2	SToRM	14.54	33.33	0.8161
	MoDL	4.62	23.81	0.3451
	Tikhonov-SToRM	16.55	35.40	0.8783
	Proposed	**20.12**	**38.97**	**0.9114**
Subject 3	SToRM	17.09	34.19	0.8345
	MoDL	4.6	21.85	0.3739
	Tikhonov-SToRM	19.02	36.17	0.8484
	Proposed	**20.73**	**37.83**	**0.9021**
Simulated dataset 1	SToRM	15.04	34.60	0.6880
	MoDL	3.72	23.27	0.4851
	Tikhonov-SToRM	15.29	34.85	0.6968
	Proposed	**23.43**	**42.98**	**0.9602**
Simulated dataset 2	SToRM	21.63	36.03	0.8641
	MoDL	4.78	19.19	0.5455
	Tikhonov-SToRM	21.92	36.92	0.8696
	Proposed	**26.82**	**41.23**	**0.9721**

Table S1: Elaboration of Table 1 in main manuscript. We compare the reconstruction methods for all test subjects across different recovery metrics. These metrics are reported for the entire field of view. Whereas, the SER (dB) metric in the previous three figures are reported for the myocardium area.

References

1 Aggarwal HK, Mani MP, and Jacob M. Modl: Model based deep learning architecture for inverse problems. IEEE Transactions on Medical Imaging, 2018; pages 1–1. ISSN 0278-0062. 10.1109/TMI.2018.2865356.