**SpezialForschungsBereich F 32**

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

# Infimal Convolution of Total Generalized Variation Functionals for dynamic MRI

M. Schloegl M. Holler A. Schwarzl

K. Bredies R. Stollberger

SFB-Report No. 2016-002 June 2016

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

Der Wissenschaftsfonds.

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

## ABSTRACT

Purpose: To accelerate dynamic MR applications using Infimal Convolution of Total Generalized Variation functionals (ICTGV) as spatio-temporal regularization for image reconstruction.

Theory and Methods: ICTGV comprises a new image prior tailored to dynamic data that achieves regularization via optimal local balancing between spatial and temporal regularity. Here it is applied for the first time to the reconstruction of dynamic MRI data. CINE and perfusion scans were investigated to study the influence of time dependent morphology and temporal contrast changes. ICTGV regularized reconstruction from subsampled MR data is formulated as a convex optimization problem. Global solutions are obtained by employing a duality based non-smooth optimization algorithm.

Results: The reconstruction error remains on a low level with acceleration factors up to 16 for both CINE and DCE MRI data. The GPU implementation of the algorithm suites clinical demands by reducing reconstruction times of one dataset to less than 4 minutes.

Conclusion: ICTGV based dMRI reconstruction allows for vast undersampling and therefore enables for very high spatial and temporal resolutions, spatial coverage and reduced scan time. With the proposed distinction of model and regularization parameters it offers a new and robust method of flexible decomposition into components with different degrees of temporal regularity.


Keywords: Dynamic MRI, CMR, Perfusion Imaging, Total Generalized Variation, Infimal Convolution, Variational Models

## INTRODUCTION

Dynamic Magnetic Resonance Imaging (dMRI) is an important technique for a broad range of applications from cardiac CINE imaging to dynamic contrast-enhanced (DCE) MRI. It generally aims to visualize time dependent changes of morphology or contrast. A common issue of all dMRI applications is the ultimate trade-off between spatial and temporal resolution. Consequently, dMRI applications benefited strongly from parallel imaging (PI) (1, 2, 3) due to the possibility of reducing the number of acquired k-space data. Successful modifications of PI techniques tailored to dMRI were reported in the past (4, 5, 6, 7).

Compressed Sensing (CS), with the mathematical foundation established in (8, 9), further improved dMRI by adding sparsity constraints to the reconstruction problem. Since the image sequence is not sparse itself, different sparsifying transformations such as a temporal Fourier-Transformation (tFT) (10, 11), a Wavelet transform (12), spatio-temporal finite differences (13) or combinations of these (14, 15) were proposed.

While, e.g., tFP turned out to be a suitable choice for periodic applications such as cardiac CINE imaging, limitations were reported for the non-periodic setting such as DCE applications. In (16), Liang proposed to reconstruct undersampled dynamic data by exploiting a representation with a low number of temporal basis functions learned by singular value decomposition from fully sampled low resolution data. This method was the basis for following work as described in (17, 18), employing principal component analysis (PCA),

and more recently in (19), to efficiently reconstruct dynamic data. Using PCA as sparsifying transform as surrogate of tFT was also discussed in (10, 15).

The exploration of data driven sparsifying transforms as with PCA assumes that the dynamic data lies in a low dimensional subspace, e.g., that an appropriate matrix representation of the image sequence is of low rank. Mathematical foundations of low-rank matrix recovery from undersampled data were studied in (20), where the nuclear norm is used as convex relaxation of the non-convex low-rank constraint. One of the first applications of such a regularization to dMRI reconstruction was reported in (21), where also the connection to the earlier work of (16) was pointed out. In consequence to former findings of using sparsity constraints for dMRI reconstructions, a joint combination of sparsity and low-rank constraints on dMRI data was developed in (22, 23, 24). In (22), non-convex Schatten p-quasi-norms were suggested as improvement to the nuclear norm, which, due to non-convexity, impose more difficulties to computing a robust numerical solution. In (25, 26) a decomposition into low-rank and sparse component (L+S) was proposed which is also referred to as robust PCA. This was applied in (27, 28, 29) to dMRI reconstruction, explicitly in combination with parallel imaging in (28). In (28), the authors argue that L+S is a natural choice for dynamic data as it models innovations on top of a highly correlated background and is therefore applicable for all kind of dMRI applications. Actually, these models require a *global* and *explicit* separation of dynamic data and background. Such a separation is, however, not generally applicable and current works employ a patch-based decomposition to *locally* improve reconstruction quality (30, 31, 32) with increased computational complexity. In (32) for instance, patches are generated from an initial dynamic CS reconstruction via motion tracking, and a patch-based low-rank reconstruction is then carried out for further improvement. A similar yet more specific approach for a compensation of breathing-motion in cardiac perfusion with motion tracking MRI was presented in (33). In (19), PCA is applied to obtain temporal basis functions from low-resolution data. These are used as $L^1$-penalized model-consistency constraints rather than imposing strict low-rank assumptions, while the model order needs to be determined heuristically. Overall, the previously discussed CS methods require an incoherence in the data domain, which can be achieved for MRI by specific modified sampling trajectories for Cartesian (34) or Non-Cartesian (35, 36) data.

Generalizing CS approaches, variational models for image reconstruction allow to introduce appropriate assumptions on an unknown object which do not necessarily rely on sparsity. Prior knowledge enters in a regularization functional which is weighted against data fidelity. For morphological image reconstruction, the Total Generalized Variation (TGV) functional (37, 38), a higher order extension of the well-known Total Variation (TV) functional (39), has been proven to be a favorable image prior in MRI reconstruction (40). By optimally balancing between different orders of differentiation, the TGV functional realizes an image model that enforces linear or polynomial smoothness while still allowing jumps, i.e., sharp boundaries.

While such a model proved to be effective for still images, the question of how to suitably extend such an approach to dynamic data was addressed recently in a general context (41). There, a convex spatio-temporal regularization functional is proposed that optimally balances between (higher order) spatial and temporal

regularity by employing an *infimal convolution of differently weighted TGV functionals* (ICTGV). This yields an analytically well-defined and computationally tractable regularization functional that enables the *automatic separation* of a dynamic sequence into components where either *strong spatial* or *strong temporal* regularization is beneficial.
In the present work, the ICTGV-functional is applied for the first time to the reconstruction of highly undersampled dynamic MR-data from different applications.

## THEORY

### Infimal Convolution of Total Variation Type Functionals

In order to achieve a stable reconstruction method for highly subsampled, noisy dMRI data, we employ spatio-temporal regularization in a variational context: Given a suitable spatio-temporal regularization functional $\mathcal{R}$ and a data fidelity term $\mathcal{D}$ modeling the MR data acquisition process with acquired noisy data $d$, we aim at reconstructing an image sequence $\hat{u}$ by solving

$$\hat{u} \in \arg\min_u \mathcal{D}(u, d) + \mathcal{R}(u).$$

Due to undersampling only a small subset of the dMRI data is known. This results in ill-posedness of the image reconstruction problem, and hence the regularization term $\mathcal{R}$ plays a crucial role.
For stationary morphologic images, regularization functionals that enforce piecewise smoothness, such as Total Variation (TV) (39) and Total Generalized Variation (TGV) (37, 38), have been successfully applied to the reconstruction from undersampled MRI data (40, 42).
Our goal is to suitably extend this concept to dynamic data by exploiting temporal correlation. For that purpose, the spatial and temporal dimension are combined in one data space, for which the different scaling of the temporal dimension introduces a parameter that has no physical meaning. For TV regularization, this parameter corresponds to a weighting of the temporal derivative and hence defines a trade-off between spatial and temporal regularization. Defining for example

$$\mathrm{TV}_\beta(u) = \|\nabla_\beta u\|_1, \quad \text{with} \quad |\nabla_\beta u| = \sqrt{(\mu_1 \partial_{\mathrm{x}} u)^2 + (\mu_1 \partial_{\mathrm{y}} u)^2 + (\mu_2 \partial_{\mathrm{t}} u)^2},$$

and $\beta = (\mu_1, \mu_2)$, the ratio $\frac{\mu_2}{\mu_1}$ fixes the space-time scaling for the whole data set and leads to a temporal regularization proportional to the value of $\frac{\mu_2}{\mu_1}$. As this ratio is not predefined, it can be tuned to adapt the regularization properties. For typical dMRI datasets, however, a uniform choice is difficult due to local, contradicting requirements. Different parts of a dynamic data set may require either a stronger spatial or stronger temporal regularization depending on the dynamic properties of the investigated region. The spatio-temporal regularization approach proposed in (41) fulfills both requirements by combining two

regularization functionals $\mathrm{TV}_{\beta_1}$ and $\mathrm{TV}_{\beta_2}$, with different parameter-tuples $\beta_1$ and $\beta_2$, enforcing stronger and weaker temporal regularization, respectively, as follows:

$$\mathrm{ICTV}(u) = \min_v \mathrm{TV}_{\beta_1}(u-v) + \gamma\,\mathrm{TV}_{\beta_2}(v).$$

This way of combining two convex functionals is well-known in convex analysis and is called infimal-convolution due to analogy to the convolution operation in Fourier analysis. A main advantage of this approach is that it allows for a local, adaptive decomposition of the dataset into two image components $(u-v)$ and $v$, each measured in terms of the respective regularization functional.

The infimal-convolution functional itself optimally balances both components and provides a first component $(u-v)$, with less dynamic information due to high temporal regularity, and a second component $v$, with more dynamic change. As we will see, such an optimal balancing comprises a flexible but tailored convex spatio-temporal regularization approach that benefits from the inherent temporal redundancy of dynamic data and, consequently, allows a high reconstruction quality even in situations with strong undersampling.

To avoid the introduction of staircasing artifacts of TV regularization, we follow the theory provided in (41) and balance between two differently weighted second-order spatio-temporal TGV functionals. As different weightings of the temporal derivative can, in a more abstract setting, be realized with a weighted gradient and weighted matrix valued symmetrized derivative, we define the ICTGV regularization functional as

$$\mathrm{ICTGV}^2_{\beta,\gamma}(u) = \min_v \mathrm{TGV}^2_{\beta_1}(u-v) + \gamma\,\mathrm{TGV}^2_{\beta_2}(v), \qquad [1]$$

where

$$\mathrm{TGV}^2_{\beta_i}(u) = \min_w \alpha_1\|\nabla_{\beta_i}u - w\|_1 + \alpha_0\|\mathcal{E}_{\beta_i}w\|_1$$

employs the weighted gradient $\nabla_{\beta_i}$ and the weighted symmetrized gradient $\mathcal{E}_{\beta_i}w = \frac{1}{2}(\nabla_{\beta_i}w + \nabla_{\beta_i}w^T)$, respectively. In practice, these weightings are chosen such that the functionals focus either on spatial or temporal regularity.

The proposed regularization functional hence optimally decomposes the image sequence in two components, both following a spatio-temporal TGV model, but with differently parametrized dynamics. It will not only favor image sequences that are sparse with respect to either of the two models, but also the ones that can be split as sum of two components, either component being preferred by one of the two parametrized TGV terms.

It was shown in (41) that ICTGV yields an analytically well defined convex regularization functional that, in the setting of differently weighting temporal and spatial derivatives, is invariant with respect to translations and rotations in space. Numerically, the solution of ICTGV regularized inverse problems can be carried out with the same techniques as for TV regularization, only the number of variables increases. In particular, in case of a convex data fidelity such as the one resulting from MR modeling, state of the art duality based optimization algorithms can be applied easily and the approximation of globally optimal solutions can be assured.

## The application of ICTGV to dynamic MRI

The application of ICTGV regularization for reconstruction of accelerated dynamic MRI amounts to solve the following minimization problem

$$\hat{u} \in \arg\min_{u} \frac{\lambda}{2} \sum_{f,c} \|\mathcal{F}_f(b_c u_f) - d_{f,c}\|_2^2 + \ \mathrm{ICTGV}^2_{\beta,\gamma}(u), \qquad [2]$$

where $f$ enumerates different time frames, $c$ the receiver coils and $\mathcal{F}_f$ denotes the Fourier operator that accounts for the acquired time dependent samples of the desired magnetization $u_f$ at frame $f$, considering the influence of the coil sensitivity distribution $b_c$. For non-Cartesian acquisitions a nonuniform Fourier-transform (43) is applied. For the solution of Eq. 2 the coil sensitivities are assumed to be known. In this work they were determined a-priori from the subsampled dataset with an algorithm described in (44). The undersampled measurements of each single coil at frame $f$ are denoted with $d_{f,c}$. Due to normal-distributed noise characteristics in MRI, the data-fidelity term is the $L^2$ discrepancy. The multiplication of the time frames $(u_f)_f$ with the sensitivity profiles $(b_c)_c$ is defined pointwise in space. The minimization problem of Eq. 2 is non-smooth but convex and can be solved by a first order duality based method (45) that results in a rather simple but efficient implementation (see Section "Discretization and Numerical Solution" in the Appendix).

## Model and regularization parameters

Variational methods use a *regularization parameter*, denoted by $\lambda$ in Eq. 2, that weights data fidelity versus regularization and should be chosen according to the expected noise level. In addition to that, the proposed regularization functional encompasses several weighting parameters that define a model of the expected image structure (*model parameters*). As described in detail below, we parametrize the ICTGV functional using *three model parameters*. This is carried out in a way that a variation of the model parameters does not influence the overall trade-off between regularization and data fidelity. Hence the three model parameters can be learned only once for an expected image structure, e.g., CINE cardiac imaging, and remain afterwards unaffected from the noise level and subsampling ratio of subsequent measurements. As automatic choice of the *regularization parameter*, we propose a linear adaption to the subsampling factor.

**Model parameters:** A first parameter that is inherent to TGV regularization is the ratio $\frac{\alpha_1}{\alpha_0}$ that defines the weighting of the different orders of differentiation. Previous TGV imaging applications and studies found that fixing this ratio to $\frac{1}{\sqrt{2}}$ yields a robust choice (40, 46), which is hence used throughout this work.
Now rewriting ICTGV as

$$\mathrm{ICTGV}_{\beta,\gamma}(u) = \min_{v} \gamma_1(s)\,\mathrm{TGV}^2_{\beta(t_1)}(u-v) + \gamma_2(s)\,\mathrm{TGV}^2_{\beta(t_2)}(v) \qquad [3]$$

lets the functional depend on three parameters:: $s \in [0,1]$ which defines the weights $\gamma_i(s)$ of the first and the second TGV functional and the $t_1$, $t_2 \in (0,\infty)$ which define the tuples $\beta(t_1)$, $\beta(t_2)$ that correspond to two

different weightings of the temporal versus the spatial derivative as detailed below. The weights $\gamma_1(s)$ and $\gamma_2(s)$ are chosen with the intention of not influencing the trade-off between data fidelity and regularization as follows:

$$\gamma_1(s) := \frac{s}{\min(s, 1-s)}, \quad \gamma_2(s) := \frac{1-s}{\min(s, 1-s)} \ .$$

In contrast to a straightforward convex combination as $s$ and $(1-s)$, this ensures that the overall cost of the functional does not reduce to zero as $s$ comes close to zero or one. The smaller of the two factors $\gamma_1(s)$ and $\gamma_2(s)$ always equals one and hence $s$ allows to balance between the $\mathrm{TGV}^2_{\beta(t_i)}$ without reducing or increasing the overall cost of $\mathrm{ICTGV}_{\beta,\gamma}$.

In order to normalize the weighting between spatial and temporal regularization we pursue the following approach: Given weights $\mu_1$ and $\mu_2$ of the spatial and temporal derivative, respectively, we require that integrating the resulting vector norm on the gradient over all possible directions gives the same results as the standard Euclidean norm. This amounts to fix

$$1 = \frac{1}{2\pi} \int_0^{\pi} \int_0^{2\pi} \sqrt{(\mu_1 \sin\theta \cos\varphi)^2 + (\mu_1 \sin\theta \sin\varphi)^2 + (\mu_2 \cos\theta)^2} \, \mathrm{d}\varphi \, \mathrm{d}\theta$$

and provides the spatial and temporal weight as function of their ratio, i.e., $(\mu_1, \mu_2) = \beta(t)$, with $t = \frac{\mu_2}{\mu_1} \in (0, \infty)$ a given ratio.

**Regularization parameter:** Given a fixed set of model parameters, our method requires only the choice of one regularization parameter. This choice should depend on the signal-to-noise-ratio (SNR), the number of voxels and time frames and the subsampling rate of the MR data. Indeed, concerning the latter, the cost of the data fidelity term decreases linearly with the subsampling ratio, since the forward operator maps only to sampled data. Based on this observation we define the regularization parameter as a linear function of the subsampling factor $r$, i.e., $\lambda(r) = kr + d$, and hence need to fix $k$ and $d$. A further adaption to the noise level and image dimension is possible but omitted for the sake of simplicity.

## METHODS

### Data acquisition and evaluation

The proposed reconstruction technique was applied to CINE cardiac MRI and to non-stress cardiac perfusion imaging. CINE cardiac MRI was evaluated with two datasets. One dataset was made available from the ISMRM reconstruction challenge 2013-2014[1] and the second dataset was a retrospectively gated bSSFP short-axis view of a health volunteer measured on 3T scanner (Skyra, Siemens Healthcare, Erlangen, Germany) with the following scan parameters: FOV = 274.62 mm × 340 mm, matrix size = 208 × 168, 25 cardiac phases with a temporal resolution of 42.72 ms, TR/TE/FA = 3.56 ms/1.78 ms/40°, 6 mm slice-thickness, TA=16 s. For reconstruction 30 channels from spine- and body-coil elements were automatically selected.

[1] `http://challenge.ismrm.org/`

One dynamic perfusion dataset was made available from New York School of Medicine (NYU). It was acquired in a healthy adult volunteer with a modified Turbo-FLASH pulse sequence on a whole-body 3T scanner (Tim Trio, Siemens Healthcare, Erlangen, Germany) using a 12-element matrix coil array and was already used for evaluation purposes of undersampled dynamic MRI reconstruction in (28).

Different levels of acceleration from 4 to 16 were simulated from the fully-sampled multi-coil data with an optimized Cartesian sampling scheme for dynamic time-series as proposed in (34), which was found to be more efficient as variable density sampling. Multi-coil reconstructions from ICTGV, spatio-temporal TGV, TV and L+S (28) were evaluated quantitatively by means of structural-similarity-index (SSIM)(47) and signal-to-error ratio (SER) as used in (22, 29), against the sum-of-squares (SOS) reconstruction of the fully sampled data within a region of interest encompassing the heart. For additional evaluation of single-coil reconstructions with kt-RPCA (29), kt-SLR (22) and kt-FOCUSS (10) the coil-combined fully sampled references were transformed to data space. Artificial degrees of subsampling from 2 to 12 were computed again with the sampling pattern described in (34) with an additional block of 8 lines of low-resolution data around the k-space-center. This results in effective acceleration factors $r \in \{1.96, 3.6, 5.2, 6.5, 7.5\}$ for the CINE cardiac and $r \in \{1.9, 3.5, 4.8, 5.7, 6.5, 7.3\}$ for the cardiac perfusion dataset.

The image reconstruction pipeline for ICTGV, TGV and TV was implemented in C++ with CUDA support using the AGILE library (48). The reconstruction framework with examples can be found online at `https://github.com/IMTtugraz/AVIONIC`. For the L+S reconstruction, the implementation provided at `http://cai2r.net/resources/software` was used, where the proposed temporal Fourier transform was replaced by a temporal TV regularization, also provided by the authors, since this improved results. For comparison to kt-RPCA, kt-SLR and kt-FOCUSS the reconstruction code provided online by the authors was used[2].

For ICTGV, a parameter learning, as described in more detail below, was carried out a-priori and parameters where then fixed for CINE and perfusion imaging for all further experiments. A similar parameter learning was also carried out for TGV and TV based reconstruction. In contrast to that, to ensure the best results, the parameters for L+S reconstruction were optimized separately for *each individual case and acceleration factor used in the evaluation*. Optimization was carried out on the parameter range as described in (28) and parameters were adapted for each case based on quantitative comparison to the (in practice unknown) ground truth. The parameter choices for kt-RPCA, kt-SLR and kt-FOCUSS were also trained a-priori according to the parameter-ranges described in (29) for one acceleration factor of $r = 8$ (with additional k-space-center lines for kt-FOCUSS). As result, the parameters were fixed to $(\mu^{\text{perf}}, \rho^{\text{perf}}) = (200, 2)$, $(\mu^{\text{CINE}}, \rho^{\text{CINE}}) = (10, 2)$ for kt-RPCA, $(\alpha^{\text{perf}}, \beta^{\text{perf}}) = (\alpha^{\text{CINE}}, \beta^{\text{CINE}}) = (500, 10^{-3})$ for kt-SLR and $\lambda^{\text{perf}} = \lambda^{\text{CINE}} = 10^{-3}$ for kt-FOCUSS, respectively. The coil sensitivities for all multi-coil reconstructions were estimated with an iterative variational approach from the subsampled data as described in (44).

[2]kt-RPCA: `http://agsp.org/bt/ktrpca/`, kt-SLR: `http://user.engineering.uiowa.edu/~jcb/Software/ktslr_matlab/Software.html`, kt-FOCUSS: `http://bispl.weebly.com/k-t-focuss.html`.

## Parameter learning

The three *model parameters* $s$, $t_1$, $t_2$ as described in Section "Model and regularization parameters" were fixed by evaluating a range of parameter sets with respect to SSIM and root-mean-square error (RMSE). For CINE cardiac imaging, two test cases from the ISMRM reconstruction challenge 2013-2014, different from the ones presented in the evaluation below, were used. The training was carried out for two factors of undersampling ($r_1 = 5$, $r_2 = 10$), for which a ground truth and a good choice of regularization parameter was known. For cardiac perfusion imaging, the model parameter training was carried out with the MRXCAT perfusion phantom (49) with a selected slice in mid-ventricular short-axis view, standard settings and additional subsampling of $r = 6$, and the perfusion dataset from NYU. As a result, the model parameters were fixed as $(t_1, t_2, s) = (4, 0.5, 0.5)$ for CINE cardiac imaging and $(t_1, t_2, s) = (9, 1, 0.6423)$ for cardiac perfusion imaging. The same test cases (the ISMRM cases for CINE imaging and the phantom for perfusion imaging) were also used for regularization parameter learning. Based on the sampling-pattern described in (34), incomplete data was generated for different subsampling rates and the reconstructions for different parameter choices were compared against the ground truth. According to these experiments, final values for $k$ and $d$ were fixed to $(k, d) = (0.34, 4.57)$ for CINE cardiac and $(k, d) = (0.08, 1.56)$ for cardiac perfusion imaging. Single-coil reconstructions were carried out with the same parameters as for the multi-coil setting.
An exemplary visualization of our results for parameter tuning is displayed in Figure 1 and Figure 2, respectively. A summary of all relevant parameter choices is given in Table 1. We also remark that, in order to remove a dependency of the regularization parameter from the signal range of the image data, we carried out a normalization to the median of the highest ten percent of the time-averaged reconstruction as preprocessing step.

## Discretization and Numerical Solution

We denote by $N$ and $M$ the image space dimensions, by $T$ the number of time frames, by $U = \mathbb{C}^{N\times M\times T}$ the space of image sequences and by $C$ the number of coils. We define a forward operator mapping from $U$ to the reduced multi-channel dMRI data $d \in \mathbb{C}^{N\times M\times T\times C}$ as

$$K : u \mapsto \left(\mathcal{F}_f \left(b_c u_f\right)\right)_{f,c}, \qquad [4]$$

following the notation defined in Eq. 2. Including the given space-time weights $(\mu_{1,i}, \mu_{2,i}) = \beta_i = \beta(t_i)$ (see Section "Parameter learning") in the differential operators we also define

$$\nabla_{\beta_i} : U \to U^3 \quad \text{and} \quad \mathcal{E}_{\beta_i} : U^3 \to U^6$$

to be a spatio-temporal gradient and spatio-temporal symmetrized gradient operator, respectively. With balancing values $\gamma_i = \gamma_i(s)$, $\|\cdot\|_2$ a discrete $L^2$ norm and $\|\cdot\|_1$ a discrete $L^1$ norm employing the standard Euclidean norm pointwise, ICTGV regularized dMRI reconstruction as in Eq. 2 can be re-written as

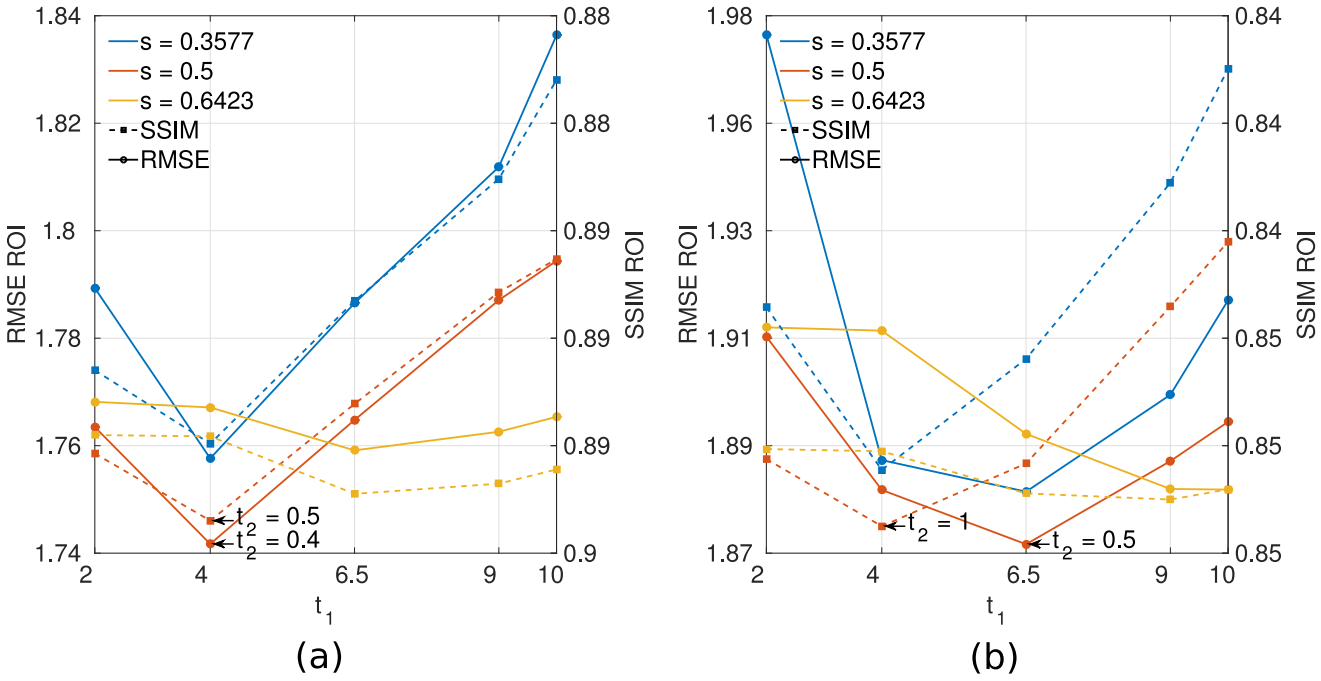



Figure 1: Evaluation of the model parameters $s, t_1, t_2$ by means of RMSE (solid lines) and SSIM (dashed lines) for a CINE cardiac test-case with acceleration factors $r = 5$ (a) and $r = 10$ (b). The colors indicate three different choices for s. The horizontal axis show different values for $t_1$ and the vertical axis the corresponding best RMSE and SSIM values achieved with $t_2 \in \{0.2, 0.4, 0.5, 1, 2.5, 3\}$, which are marked (arrow) with the corresponding best value for $t_2$.

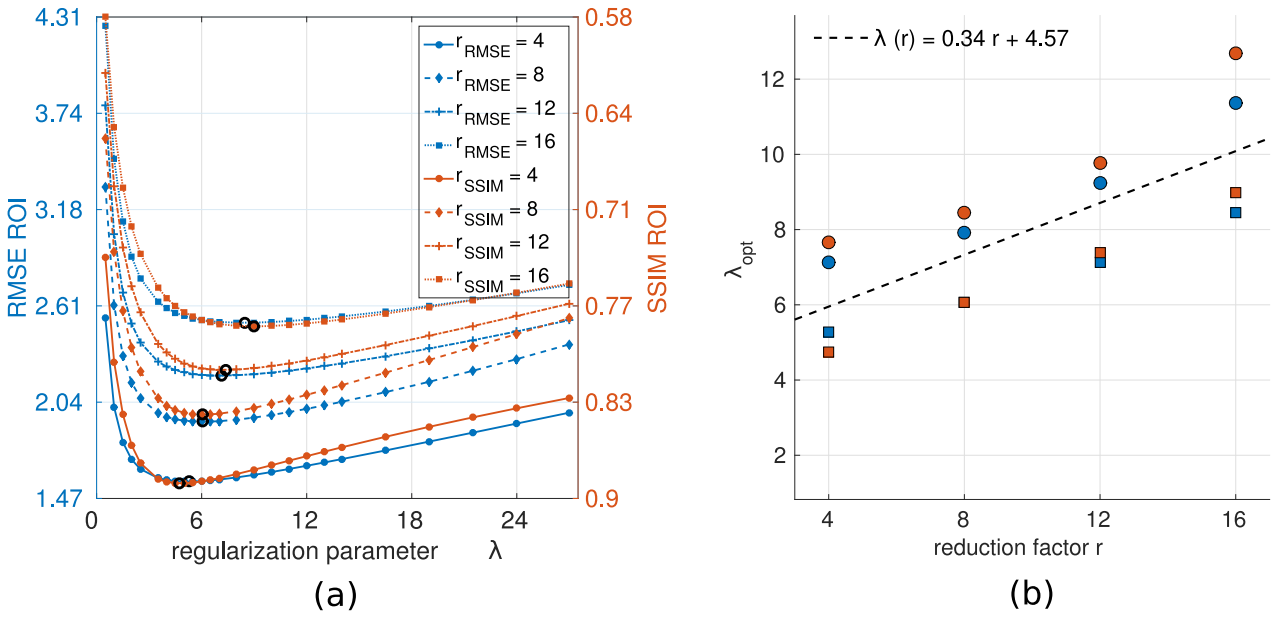



Figure 2: (a) RMSE (blue curves) and SSIM (red curves) evaluation exemplified for one CINE cardiac test case for different regularization parameters $\lambda$ and acceleration factors $r = 4$ (solid line), $r = 8$ (dashed line), $r = 12$ (dashed-dotted line) and $r = 16$ (dotted line). The corresponding optimal values for $\lambda$ are indicated with black circles and were calculated by spline-interpolation between the used sample-points. (b) Optimal values for $\lambda$ according to RMSE (blue) and SSIM (red) for different acceleration factors and the test case displayed in (a) (squares) and the second test case (dots). The linear regression to both cases and metrics (black dashed line) yields the proposed parameter choice.

Table 1: Regularization and model parameter choice

| Symbol | Denomination | Chosen value | | Testing range |
|---|---|---|---|---|
| Model parameter | | CINE | Perf. | |
| $\alpha_0/\alpha_1$ | Ratio of TGV weights | $1/\sqrt{2}$ | $1/\sqrt{2}$ | $--$ |
| $t_1 \rightarrow \beta(t_1)$ | Spatio-temporal weight 1st component | 4 | 9 | $[2, 4, 6.5, 9, 10]$ |
| $t_2 \rightarrow \beta(t_2)$ | Spatio-temporal weight 2nd component | 0.5 | 1 | $[0.2, 0.4, 0.5, 1, 2.5, 3]$ |
| $s \rightarrow \gamma_{1,2}(s)$ | Weighting between 1st and 2nd component | 0.5 | 0.6423 | $[0.3577, 0.5, 0.6423]$ |
| Regularization parameter | | | | |
| $\lambda(r) = kr + d$ | Regularization parameter | $--$ | | $\lambda \in [0.5, \cdots, 27]$ $r \in [4, 8, 12, 16]$ |
| $k$ | Slope | 0.34 | 0.08 | $--$ |
| $d$ | Intercept | 4.57 | 1.56 | $--$ |

$$\min_{u,w_1,v,w_2} \frac{\lambda}{2}\|Ku - d\|_2^2 + \gamma_1 \left(\alpha_1\|\nabla_{\beta_1}(u - v) - w_1\|_1 + \alpha_0\|\mathcal{E}_{\beta_1} w_1\|_1\right) + \gamma_2 \left(\alpha_1\|\nabla_{\beta_2} v - w_2\|_1 + \alpha_0\|\mathcal{E}_{\beta_2} w_2\|_1\right). \quad [5]$$

From an abstract viewpoint Eq. 5 constitutes a convex problem of the form

$$\min_x F(Hx)$$

with $H$ being a linear operator and $F$ a convex, lower semi-continuous functional. Employing convex duality, i.e., that $F(x) = \sup_y (x, y) - F^*(y)$ with $F^*$ denoting the convex conjugate of $F$, we can further obtain an equivalent saddle-point reformulation as

$$\min_x \max_y \ (Hx, y) - F^*(y).$$

This saddle-point problem is then solved by the primal-dual algorithm of (45) in a way that requires only simple, explicit calculations and such that global convergence to an optimal solution is guaranteed. Proper convergence of the algorithm was ensured by computing an approximation of a *duality gap* (46) that is known to be zero only for the optimal solutions. As, according to our experience, 500 iterations are sufficient for the duality gap to indicate a voxel-wise error in the range of $\left[10^{-5}, 10^{-2}\right]$, this number of iterations was used for all experiments.

Table 2: ICTGV reconstruction times in seconds (500 iterations), obtained with a MATLAB CPU implementation (Intel i5-2500K, 3.30GHz) and a CUDA GPU implementation (NVidia GeForce GTX 770), for different data-dimensions (3rd to 5th column) denoted as $(n = N \times M \times T, C)$. The top rows depict the time in seconds to perform 500 iterations while the bottom rows show the total reconstruction time including coil sensitivity estimations.

| Device | | $(6.553 \times 10^5, 12)$ | $(1.747 \times 10^5, 30)$ | $(2.328 \times 10^5, 32)$ |
|---|---|---|---|---|
| CPU | Iterations | 1325.38 | 5783.43 | 8919.39 |
| | Total | 1365.86 | 6238.18 | 9355.47 |
| GPU | Iterations | 34.54 | 134.34 | 199.33 |
| | Total | 39.73 | 176.61 | 244.61 |

### GPU implementation

In order to speed up reconstruction times, the MATLAB prototype was translated to C++/CUDA enabling highly parallel GPU computations. As a consequence, the computational speed was improved drastically with up to 50-fold acceleration. Table 2 summarizes reconstruction times for typical data dimensions and a sufficient number of iterations.

## RESULTS

Figure 3 shows ICTGV reconstructions of a CINE cardiac dataset from the ISMRM challenge for different acceleration factors. The reconstructions exhibit a high degree of congruence to the fully-sampled reconstruction with improved noise-suppression in the background up to very high acceleration factors.

Figure 4 shows down-sampling experiments for the measured short-axis dataset with an additional comparison to L+S reconstruction for acceleration factors of 12 and 16. Again the ICTGV reconstruction reaches excellent quality up to $r = 12$ with slightly reduced fidelity for $r = 16$, while L+S based reconstructions exhibit corruption with residual undersampling and temporal blocking artifacts. We highlight again, that for L+S the parameters were trained for the same test case and each acceleration factor separately, while for ICTGV parameters were fixed a-priori based on different data.

Figure 5 displays the results from a down-sampling experiment using perfusion data, where a good delineation of the myocardial wall and the papillary muscles was achieved up to $r = 16$. Reconstruction quality with L+S is slightly worse than with ICTGV, in particular a loss of spatial details is apparent.

A quantitative evaluation by means of SSIM and SER against the fully sampled reference as ground-truth is summarized in Table 3. There, ICTGV regularization is compared against spatio-temporal $\mathrm{TGV}^2_\beta$ and

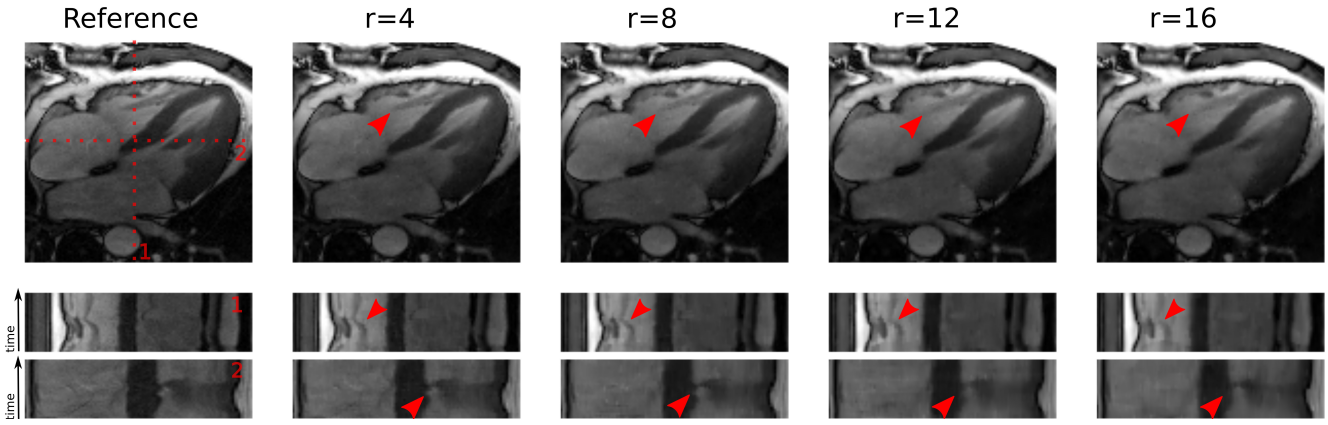



Figure 3: Magnitude images from simulated accelerations $r = (4, 8, 12, 16)$ for the four-chamber-view bSSFP dataset. The fully sampled sum-of-squares reconstruction is displayed in the 1st column and the therein indicated time-lines are shown in the 2nd and 3rd row.

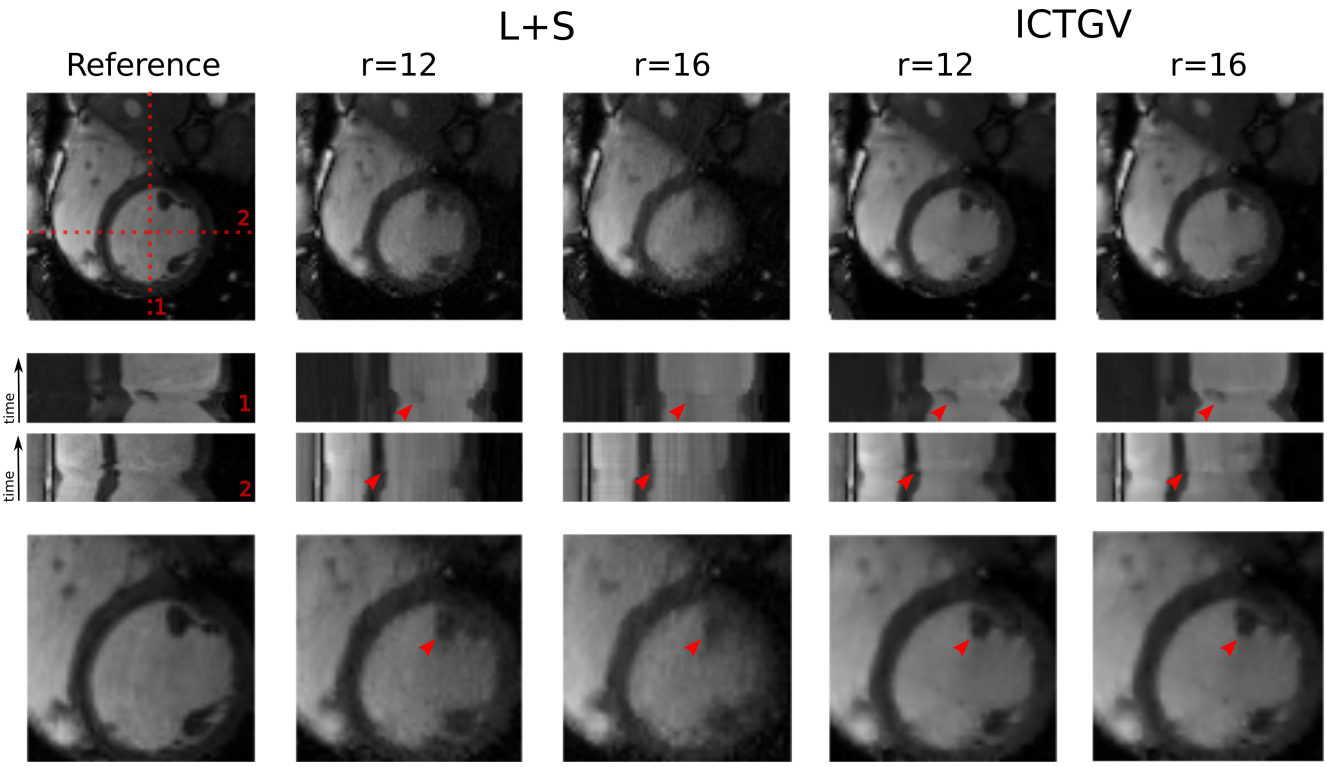



Figure 4: Comparison of magnitude images of fully sampled reference reconstruction (1st column), L+S (2nd and 3rd column) and ICTGV reconstruction (4th and 5th column) for bSSFP CINE cardiac data acquired in short-axis-view and undersampling factors of $r = (12, 16)$. A late diastolic time-frame is displayed in the first row and indicated vertical and horizontal time-lines in the 2nd and 3rd row. A closeup of the heart region is displayed in the 4th row.

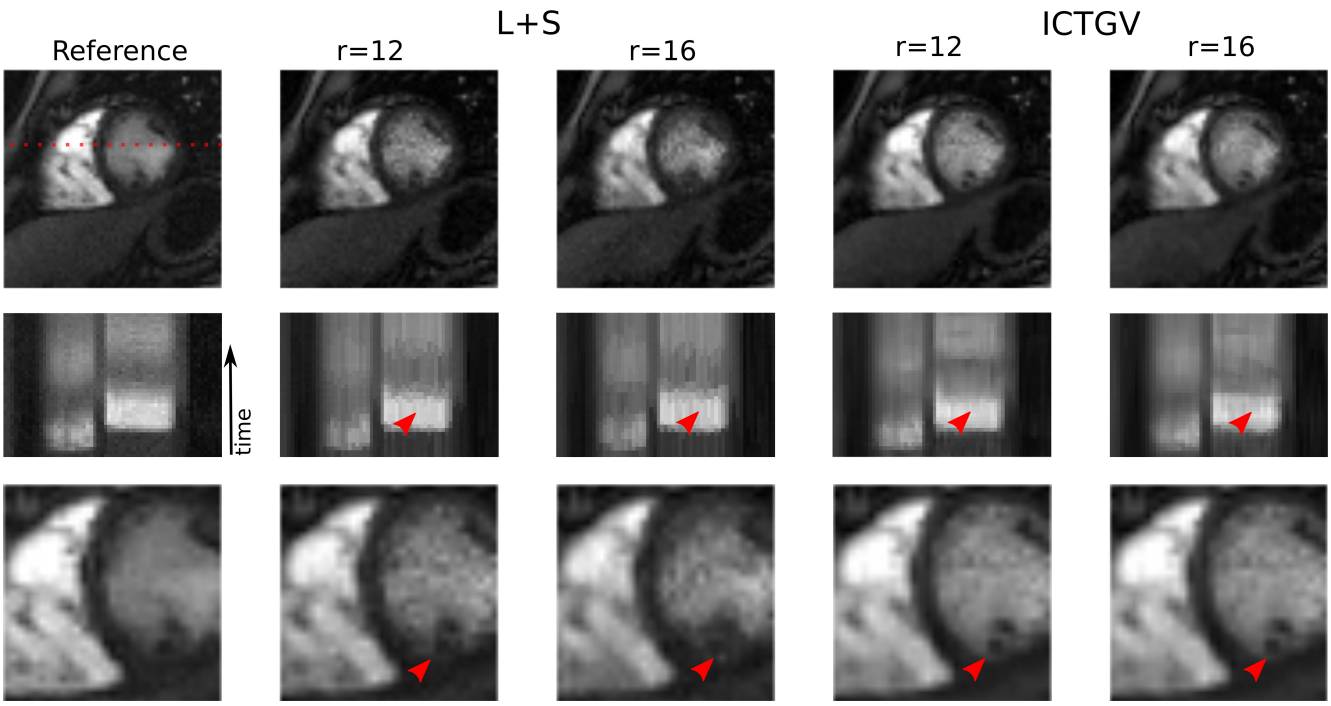



Figure 5: Comparison of magnitude images of fully sampled reference reconstruction (1st column), L+S (2nd and 3rd column) and ICTGV reconstruction (4th and 5th column) for the cardiac perfusion dataset and undersampling factors of $r = (12, 16)$. A selected time-frame is displayed in the first row and the indicated time-line in the second row. A close-up of the heart region is displayed in the third row.

$TV_\beta$ regularization and L+S reconstruction. Model and regularization-parameter training as described in Section "Theory" was also carried out both for $TGV^2_\beta$ and $TV_\beta$ reconstruction. For the L+S reconstruction parameter learning was also carried out by means of SSIM and SER for each individual test case and acceleration factor of the evaluation. For CINE cardiac test cases, ICTGV almost always scores best for both metrics with considerable improvement against L+S, in particular for higher acceleration factors. Compared to temporal $TGV^2_\beta$ and $TV_\beta$, reconstruction quality improves slightly. In contrast to that, a substantial increase is observable when comparing ICTGV reconstruction against temporal $TGV^2_\beta$ and $TV_\beta$ for the cardiac perfusion test case. Comparison to L+S for the perfusion case again shows a solid improvement with ICTGV. We also refer to Supporting Figure S1 for a visual comparison of ICTGV against TGV for multi-coil reconstruction of both, CINE cardiac imaging ($r = 16$) and cardiac perfusion ($r = 12$) reconstructions. While the results for the CINE dataset are similar, the perfusion results show a loss of spatial details and a temporal blurring with spatio-temporal TGV. In particular, the ICTGV reconstruction still allows to detect small image features that are lost with TGV.

Reconstruction results for single-coil data with ICTGV, kt-RPCA, kt-SLR and kt-FOCUSS from undersampled cardiac perfusion ($r = 7.3$) and CINE cardiac imaging ($r = 3.6$, $r = 7.5$) are displayed in Figure 6 and 7, respectively, together with the computed SER values in dB and indicated time courses (CINE) or the mean intensity within a ROI in the left ventricle (perfusion). A more detailed summary of quantitative evaluation results by means of SER and SSIM for acceleration factors as provided in the Section "Methods" is given in Supporting Table S1. The results show advantages of ICTGV in terms of error measures, consistently for all acceleration factors. For the perfusion case, the ICTGV signal time course is very close to the reference data set. The single time frames of ICTGV and kt-SLR appear somewhat denoised, and some details, as highlighted by arrows, are best visible in the ICTGV reconstruction. The individual images of all CINE reconstructions appear similar, however, the x-t plot exhibits some rippling for kt-FOCUSS and particular for kt-RPCA.

For the perfusion dataset the decomposition into ICTGV components for $r = 8$ is shown in Supporting Figure S2. Here, the static background, slower contrast dynamics with increased temporal blurring as well as morphologic changes are stored within the first component, while more rapid intensity changes (ventricles) are mapped in the second component. The separation is also displayed by the mean intensity change (magnitude) within the right ventricle (Sup. Fig. S2 b) and myocardium (Sup. Fig. S2 c), where a high agreement of the ICTGV reconstruction to the fully sampled reference is observable.

## DISCUSSION

The results presented in this work demonstrate the performance of ICTGV as new spatio-temporal regularization approach for the reconstruction of undersampled dynamic multi-coil MRI data. Two exemplary application scenarios were considered: Cardiac CINE imaging with quasi-periodic morphological motion and

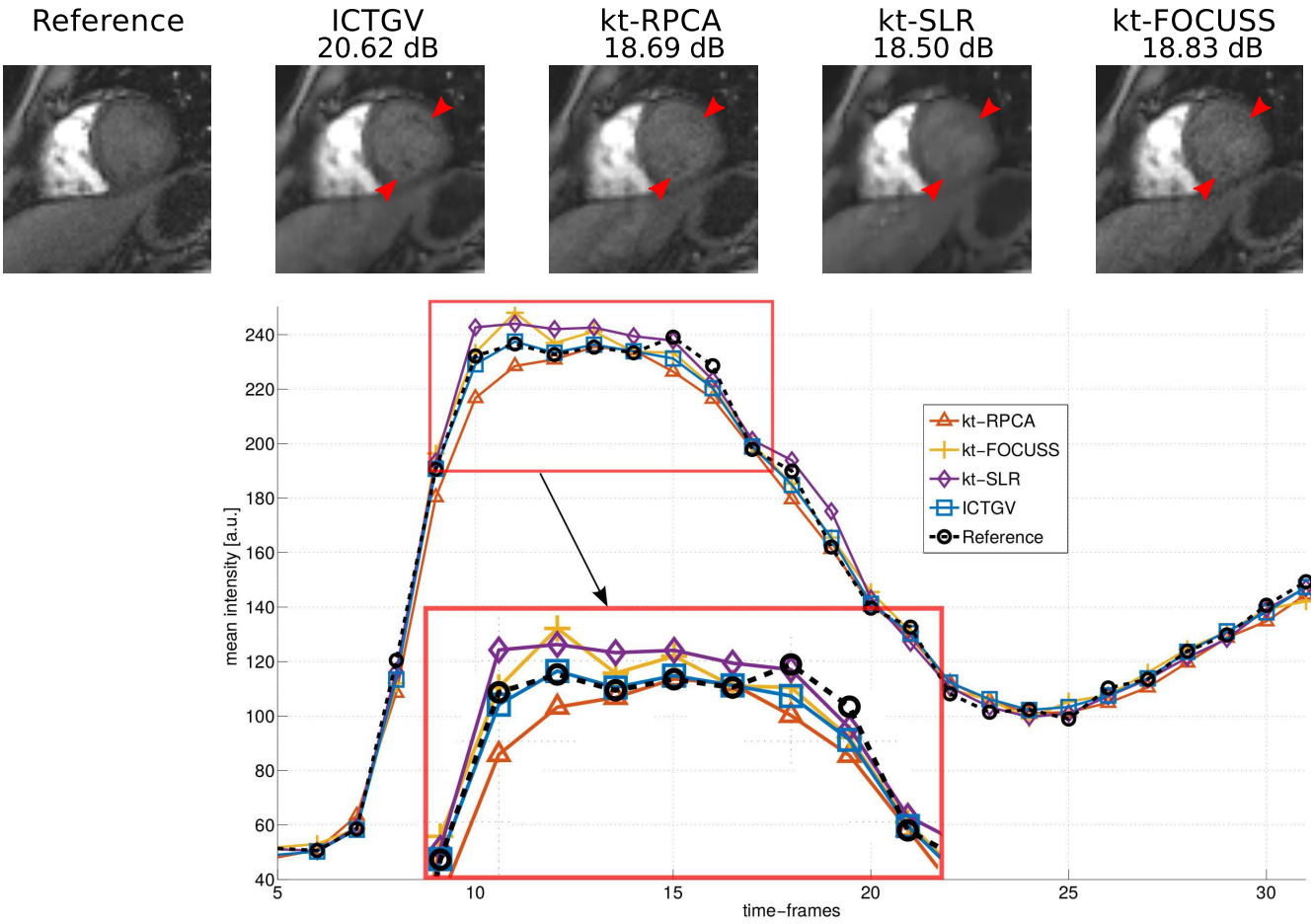



Figure 6: Coil-combined fully sampled reference (1st column), ICTGV (2nd column), kt-RPCA (3rd column), kt-SLR (fourth column) and kt-FOCUSS (fifth column) single-coil reconstructions from undersampled Cartesian data ($r = 7.3$) for the cardiac perfusion dataset with computed SER values in dB. The mean time-course of an indicated $3 \times 3$ voxel region within the right ventricle is plotted for all methods under investigation with a closeup view of the peak-signal.

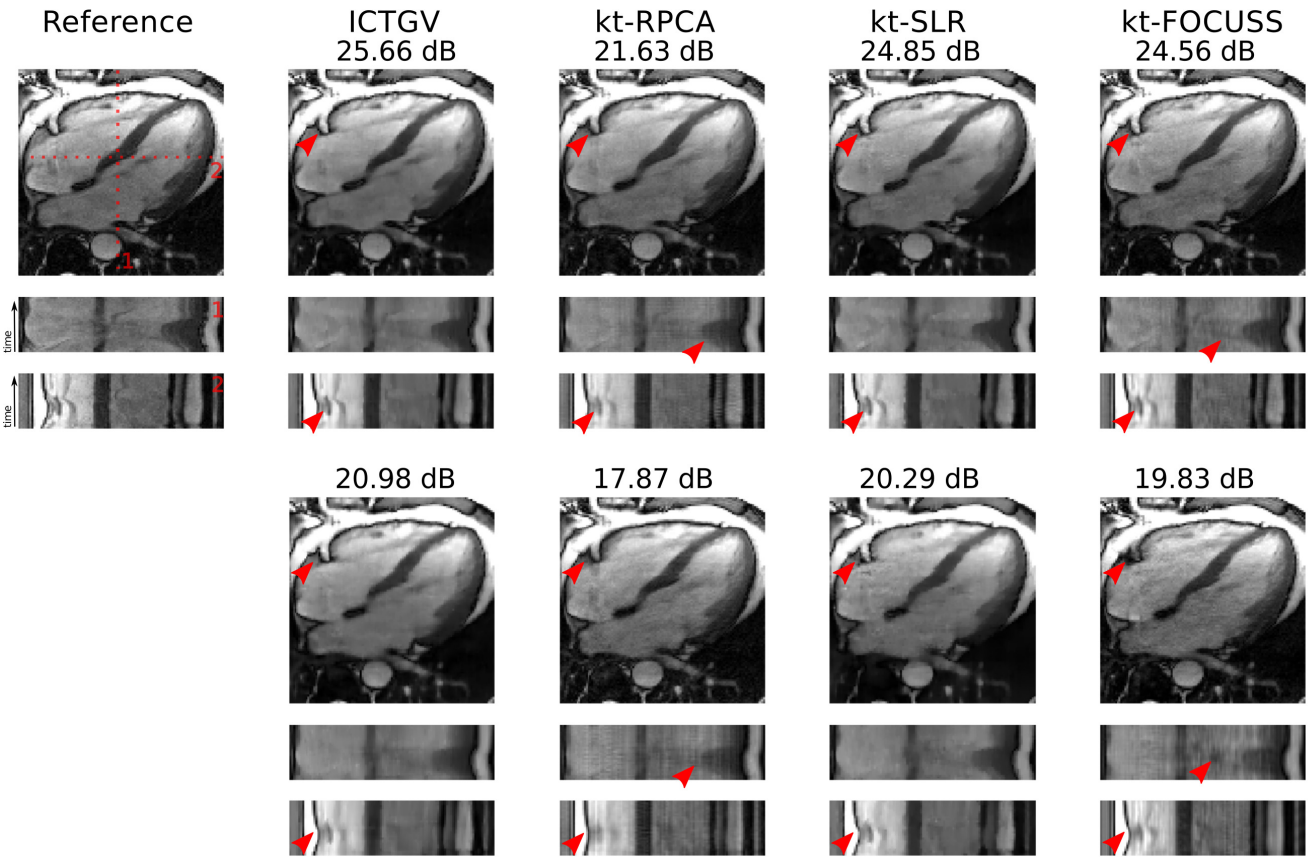



Figure 7: Coil-combined fully sampled reference (1st column), ICTGV (2nd column), kt-RPCA (3rd column), kt-SLR (fourth column) and kt-FOCUSS (fifth column) single-coil reconstructions from undersampled Cartesian data with $r = 3.6$ (1st to 3rd row) and $r = 7.5$ (4th to 6th row) for the four chamber CINE cardiac dataset. Reconstructions are displayed with a selected time-frame and the indicated horizontal and vertical time-lines.

Table 3: Quantitative evaluation (best values underlined) of multi-coil reconstruction for ICTGV against $\mathrm{TGV}^2_\beta$, $\mathrm{TV}_\beta$ and L+S reconstruction for CINE cardiac and cardiac perfusion cases by means of SSIM and SER against the fully sampled sum-of-squares reconstruction.

| | ICTGV | | $\mathrm{TGV}^2_\beta$ | | $\mathrm{TV}_\beta$ | | L+S | |
|---|---|---|---|---|---|---|---|---|
| | SER (dB) | SSIM | SER (dB) | SSIM | SER (dB) | SSIM | SER (dB) | SSIM |
| SA view | | | | | | | | |
| $r = 4$ | 25.79 | <u>0.9167</u> | <u>25.85</u> | 0.9164 | 25.60 | 0.9154 | 25.78 | 0.9154 |
| $r = 8$ | <u>23.18</u> | 0.8738 | 23.13 | <u>0.8744</u> | 22.89 | 0.8690 | 21.75 | 0.8475 |
| $r = 12$ | <u>21.45</u> | <u>0.8376</u> | 21.37 | 0.8374 | 21.21 | 0.8320 | 19.09 | 0.7802 |
| $r = 16$ | <u>20.47</u> | <u>0.8160</u> | 20.41 | 0.8147 | 20.28 | 0.8107 | 17.31 | 0.7177 |
| Four Chamber | | | | | | | | |
| $r = 4$ | 20.32 | <u>0.8267</u> | <u>20.37</u> | 0.8216 | 20.26 | 0.8220 | 20.29 | 0.8109 |
| $r = 8$ | <u>19.72</u> | <u>0.7787</u> | <u>19.72</u> | 0.7726 | 19.64 | 0.7718 | 19.54 | 0.7408 |
| $r = 12$ | <u>19.17</u> | <u>0.7445</u> | 19.16 | 0.7399 | 19.11 | 0.7397 | 18.36 | 0.6853 |
| $r = 16$ | <u>18.94</u> | <u>0.7211</u> | 18.92 | 0.7165 | 18.84 | 0.7154 | 17.43 | 0.6346 |
| Cardiac Perfusion | | | | | | | | |
| $r = 4$ | <u>21.30</u> | <u>0.8631</u> | 20.91 | 0.8554 | 19.91 | 0.8256 | 20.74 | 0.8564 |
| $r = 8$ | <u>20.34</u> | <u>0.8347</u> | 19.77 | 0.8191 | 18.16 | 0.7554 | 19.58 | 0.8172 |
| $r = 12$ | <u>18.71</u> | <u>0.8059</u> | 18.28 | 0.7817 | 16.57 | 0.6895 | 17.79 | 0.7855 |
| $r = 16$ | <u>17.54</u> | <u>0.7813</u> | 17.28 | 0.7495 | 15.38 | 0.6407 | 17.01 | 0.7597 |

cardiac perfusion imaging reflecting contrast changes to due alteration of tissue relaxation properties. The corresponding experiments show the capability of the proposed method to obtain artifact free results with high temporal fidelity and improved noise suppression for both applications up to high acceleration factors. This is confirmed both quantitatively, in terms of SSIM and SER, as well as qualitatively with selected frames and the visualization of time lines.

Furthermore, the proposed method was evaluated against state of the art reconstruction methods for single-coil and multi-coil reconstruction. In the multi-coil setting, a comparison to the L+S approach shows improved performance both for CINE and perfusion imaging. This is quantified in SER and SSIM error metrics and can also be observed visually. Interestingly, also spatio-temporal TGV and TV regularization lead to comparable or improved results compared to L+S in terms of error metrics. One possible explanation for that is that L+S in fact does not make use of any regularity in space. Spatial regularity, however, is a strong source of redundancy which is exploited both by spatio-temporal TV/TGV and ICTGV. A second explanation might be the global nature of the low-rank prior, which is not able to adapt locally to contrast or morphological changes, a shortcoming which is overcome by ICTGV regularization. The kt-RPCA (29) method is, in principle, identical with L+S, yet, in contrast to the multi-coil reconstruction with L+S, tFT is used instead of temporal TV. While L+S also proposed tFT as preferred choice, temporal TV gave improved results in our experiments and was hence used for comparison. The kt-SLR method on the other hand imposes sparsity constraints with spatio-temporal TV and non-convex low-rank constraints with Schatten p-quasi-norms ($p < 1$) jointly, instead of the proposed decomposition approach. Using non-convex quasi-norms results in a challenging optimization problem, yet, with the parameters tuning as suggested in (29), reconstruction results remain very competitive to ICTGV. Computationally, however, ICTGV reconstruction has the advantage of comprising the solution of a convex optimization problem, for which to proposed numerical algorithm guarantees global convergence. Also reconstructions gained with kt-FOCUSS, that conditionally requires low-resolution data, remain competitive for both applications yet with increased residual flickering in the temporal domain.

To assess the benefit of the proposed balancing between spatial and temporal regularization, we have also implemented straightforward spatio-temporal Total Variation and second order Total Generalized Variation regularization (itself never applied to dynamic MR reconstruction). A quantitative evaluation for different acceleration factors, test cases and error metrics shows the superiority of ICTGV, in particular for perfusion imaging (Table 3). The stronger improvement obtained with perfusion imaging can be explained by the observation that rapid intensity changes due to contrast inflow make a decomposition to different scales of temporal regularity even more beneficial. The comparison of Supporting Figure S1 further shows visual differences between ICTGV and TGV reconstructions, which are apparent for the perfusion case, where small features are lost with TGV but still recovered with ICTGV. Overall, the experiments confirm that, even though the differences with CINE imaging might be subtle, ICTGV is consistently superior to spatio-temporal TGV over different experimental setups.

In this context, it is also interesting to note that for spatio-temporal TV and TGV regularization, the parameter defining the ratio between temporal and spatial regularization, denoted by $t$, was optimized to achieve the best results. As can be seen in the plots provided in the Supporting Figure S3, the optimal value was found to be roughly in the interval [3,5] for both approaches, with a strong decrease of image quality for $t \to 0$ and $t \to \infty$. Hence a small or large choice of $t$, which for spatio-temporal TV/TGV approximately yields pure spatial and pure temporal TV/TGV regularization, respectively, significantly worsens reconstruction quality. This indicates that both, pure spatial and pure temporal TV/TGV regularization are not sufficient to achieve state-of-the art results for undersampled dMRI reconstruction and a combined exploration of spatio-temporal redundancies is necessary.

An additional feature of the presented method is that a decomposition into two components is obtained. For cardiac perfusion imaging, slower portions of the intensity changes due to the passing contrast agent, e.g., in the myocardium, as well as changes in morphology are accumulated in the first component, while regions of fast intensity changes within the ventricles, liver and kidney are captured by the second component (see, again, Sup. Fig. S2). This has similarities to a decomposition into a temporal component with correlated background (low-rank) and another with temporal changes (sparse). Yet our methods acts *locally*, while the low-rank assumption is inherently global. The local distribution of the image content to the two ICTGV components depends on the model parameters, which were optimized for the overall reconstruction quality. An exploration of ICTGV reconstruction for applications that require a different weighting, e.g., time-resolved MR angiography, where the interest is not the full image, but components with specific contrast dynamics, will be subject to further research.

While the present work proposes the infimal-convolution of two second order TGV functionals for spatio-temporal regularization, the analytical framework of ICTGV as presented in (41) as well as our numerical implementation, in principle, allow the inclusion of arbitrary many TGV functionals for balancing, also using different, higher orders of differentiation. This might be of interest in cases where one aims at resolving different types or scales of motion, also beyond the spatio-temporal regime. An example would be MR parameter mapping and MR fingerprinting where, in the spirit of (50, 51), one could use higher-order differentiation for regularization along the parameter dimension.

The presented method clearly distinguishes between model and regularization parameters. The assumption that the former influence the image model but are independent of the overall trade-off between regularization and data fidelity has been confirmed by experiments showing that the optimal choice of model parameters is robust along different subsampling rates. For the choice of regularization parameters, the proposed linear adaptation constitutes a heuristic to compensate for alteration of the data-fidelity-cost due to subsampling which was confirmed to be effective along a long range of subsampling rates by experimental results. Furthermore, these experiments showed robustness of our method with respect to the regularization parameter choice in the sense that reconstruction quality decreases only slightly for parameter choices within a range of roughly $\lambda = \lambda_{opt} \pm 10\%$ from the optimum (see Figure 2). We point out that the proposed parametrization strategies

might be of general interest for other iterative reconstruction approaches, since most state-of-the-art methods depend on different model parameters.

Finally, ICTGV-regularized dMRI reconstruction poses a standard non-smooth convex optimization problem that can efficiently be solved with the primal-dual algorithm proposed in (45). Due to convexity, it is possible to guarantee convergence to a global optimum and also to quantify the solution error via a primal-dual gap. The GPU based framework that was developed for this approach performs the whole reconstruction pipeline within a few minutes and is able to handle Cartesian and Non-Cartesian trajectories and the vendor independent ISMRMRD data-format (52). It is available online at `https://github.com/IMTtugraz/AVIONIC`.

## CONCLUSIONS

The proposed ICTGV-based method constitutes a robust reconstruction framework for highly accelerated dMRI. Our experiments confirm a good visual representation of morphological details as well as contrast dynamics for acceleration factors of 12 and beyond. Consequently, our method addresses clinical demands of reduced scan times, higher spatial and temporal resolutions and spatial coverage, while it mitigates problems from incomplete breath-hold capabilities or patient compliance. The developed algorithm is able to incorporate promising developments on non-Cartesian sequence design (53, 54). The provided GPU accelerated implementation can process the manufacturer independent ISMRMRD data standard completing the whole reconstruction framework within a few minutes and thereby it enables the applicability of the proposed method in clinical practice. As an additional feature, the method allows a local separation of components beyond the paradigm of background and dynamic information and provides a model of different temporal scales of motion in MRI – a potential that is yet to be explored for specific applications. A future goal will be to extend our approach to volumetric data. While conceptually the method is not limited to any particular space dimension, such an extension will pose additional numerical challenges in terms of computation time and memory requirement. To maintain practical applicability, more sophisticated algorithms such as (55, 56) might hence be necessary.

## ACKNOWLEDGMENTS

The authors would like to thank Ricardo Otazo from NYU for providing the cardiac perfusion data, Alexey Samsonov and Sebastian Kozerke for providing the fully sampled data from the ISMRM 2013-2014 reconstruction challenge and Gert and Ursula Reiter from Medical University of Graz for help with data collection and support for evaluation. This work is funded and supported by the Austrian Science Fund (FWF) under grant “SFB F32" (SFB “Mathematical Optimization and Applications in Biomedical Sciences”).

## APPENDIX: NUMERICAL SOLUTION

As described in Section "Discretization and Numerical Solution", it is our goal to reconstruct an image sequence $u \in U$ by solving

$$\min_{u,w_1,v,w_2} \frac{\lambda}{2}\|Ku-d\|_2^2 + \gamma_1\left(\alpha_1\|\nabla_{\beta_1}(u-v)-w_1\|_1 + \alpha_0\|\mathcal{E}_{\beta_1}w_1\|_1\right) + \gamma_2\left(\alpha_1\|\nabla_{\beta_2}v-w_2\|_1 + \alpha_0\|\mathcal{E}_{\beta_2}w_2\|_1\right),$$

where $\nabla_{\beta_i} : U \to U^3$ and $\mathcal{E}_{\beta_i} : U^3 \to U^6$ are defined as

$$\nabla_{\beta_i} u = (\mu_{1,i}\delta_{\mathrm{x}+}u, \mu_{1,i}\delta_{\mathrm{y}+}u, \mu_{2,i}\delta_{\mathrm{t}+}u)$$

and

$$\mathcal{E}_{\beta_i} w = \left(\mu_{1,i}\delta_{\mathrm{x}-}w^1, \mu_{1,i}\delta_{\mathrm{y}-}w^2, \mu_{2,i}\delta_{\mathrm{t}-}w^3, \frac{\mu_{1,i}\delta_{\mathrm{y}-}w^1 + \mu_{1,i}\delta_{\mathrm{x}-}w^2}{2}, \frac{\mu_{2,i}\delta_{\mathrm{t}-}w^1 + \mu_{1,i}\delta_{\mathrm{x}-}w^3}{2}, \frac{\mu_{2,i}\delta_{\mathrm{t}-}w^2 + \mu_{1,i}\delta_{\mathrm{y}-}w^3}{2}\right).$$

The operators $\delta_{\mathrm{x}+}$, $\delta_{\mathrm{y}+}$, $\delta_{\mathrm{t}+}$ and $\delta_{\mathrm{x}-}$, $\delta_{\mathrm{y}-}$, $\delta_{\mathrm{t}-}$ define symmetrically extended forward and backward finite difference operators, respectively, with respect to the x, y and t coordinate.

The $L^2$ norm is defined for $d \in \mathbb{C}^{N\times M\times T\times C}$ as

$$\|d\|_2^2 = \sum_{i,j,f,c} |d_{i,j,f,c}|^2$$

and, abusing notation, the norm $\|\cdot\|_1$ is defined for $w = (w^1, w^2, w^3) \in U^3$ as

$$\|w\|_1 = \sum_{i,j,f} \sqrt{|w^1_{i,j,f}|^2 + |w^2_{i,j,f}|^2 + |w^3_{i,j,f}|^2}$$

and for $\xi = (\xi^1, \xi^2, \xi^3, \xi^4, \xi^5, \xi^6) \in U^6$ as

$$\|\xi\|_1 = \sum_{i,j,f} \sqrt{|\xi^1_{i,j,f}|^2 + |\xi^2_{i,j,f}|^2 + |\xi^3_{i,j,f}|^2 + 2|\xi^4_{i,j,f}|^2 + 2|\xi^5_{i,j,f}|^2 + 2|\xi^6_{i,j,f}|^2},$$

where the factor 2 in front of $\xi_4$, $\xi_5$, $\xi_6$ compensates for the symmetrization of the Jacobian in the definition of $\mathcal{E}_{\beta_i}$.

It is our goal to obtain a saddle-point problem of the form

$$\min_x \max_y \ (Hx, y) - F^*(y)$$

that is equivalent to our original problem. To this aim, first note that ICTGV can be reformulated as

$$\mathrm{ICTGV}^2_{\beta,\gamma}(u) = \min_{x=(u,w_1,v,w_2)} \|H_1 x\|_{1,\alpha,\gamma}$$

with

$$H_1 = \begin{pmatrix} \nabla_{\beta_1} & -\mathrm{Id} & -\nabla_{\beta_1} & 0 \\ 0 & \mathcal{E}_{\beta_1} & 0 & 0 \\ 0 & 0 & \nabla_{\beta_2} & -\mathrm{Id} \\ 0 & 0 & 0 & \mathcal{E}_{\beta_2} \end{pmatrix}, \qquad [6]$$

and

$$\|(p_1, q_1, p_2, q_2)\|_{1,\alpha,\gamma} = \gamma_1 \left(\alpha_1 \|p_1\|_1 + \alpha_0 \|q_1\|_1\right) + \gamma_2 \left(\alpha_1 \|p_2\|_1 + \alpha_0 \|q_2\|_1\right)$$

representing a weighted $L^1$ norm. We further need the convex conjugates of $\|\cdot\|_{1,\alpha,\gamma}$ and $\frac{\lambda}{2}\|\cdot - d\|_2^2$ which are given for $z = (p_1, q_1, p_2, q_2)$ as

$$\|z\|_{1,\alpha,\gamma}^* := \sup_{z'} (z, z') - \|z'\|_{1,\alpha,\gamma} = \mathcal{I}_{\{\|\cdot\|_{\infty,\alpha,\gamma} \leq 1\}}(z),$$

where

$$\mathcal{I}_{\{\|\cdot\|_{\infty,\alpha,\gamma} \leq 1\}}(z) = \begin{cases} 0 & \text{if } \max\{\gamma_1\alpha_1\|p_1\|_\infty, \gamma_1\alpha_0\|q_1\|_\infty, \gamma_2\alpha_1\|p_2\|_\infty, \gamma_2\alpha_0\|q_2\|_\infty\} \leq 1, \\ \infty & \text{else,} \end{cases}$$

and

$$(\frac{\lambda}{2}\|\cdot - d\|_2^2)^*(r) := \sup_{r'} (r', r) - \frac{\lambda}{2}\|r' - d\|_2^2 = \frac{1}{2\lambda}\|r\|_2^2 + (d, r).$$

Then, a formulation equivalent with the minimization problem in Eq. 2 is obtained as

$$\begin{aligned} \min_u \frac{\lambda}{2}\|Ku - d\|_2^2 + \mathrm{ICTGV}_{\beta,\gamma}^2(u) &\Leftrightarrow \min_{x=(u,w_1,v,w_2)} \frac{\lambda}{2}\|Ku - d\|_2^2 + \|H_1 x\|_{1,\alpha,\gamma} \\ &\Leftrightarrow \min_{x=(u,w_1,v,w_2)} \max_{y=(z,r)} (Ku, r) - (d, r) - \frac{1}{2\lambda}\|r\|_2^2 + (H_1 x, z) - \mathcal{I}_{\{\|\cdot\|_{\infty,\alpha,\gamma} \leq 1\}}(z) \\ &\Leftrightarrow \min_{x=(u,w_1,v,w_2)} \max_{y=(z,r)} (Hx, y) - (d, r) - \frac{1}{2\lambda}\|r\|_2^2 - \mathcal{I}_{\{\|\cdot\|_{\infty,\alpha,\gamma} \leq 1\}}(z) \\ &\Leftrightarrow \min_{x=(u,w_1,v,w_2)} \max_{y=(z,r)} (Hx, y) - F^*(y). \end{aligned}$$

with $H = \begin{pmatrix} H_1 \\ K_1 \end{pmatrix}$, $K_1 x = Ku$, and

$$F^*(y) = F^*(z, r) = (d, r) + \frac{1}{2\lambda}\|r\|_2^2 + \mathcal{I}_{\{\|\cdot\|_{\infty,\alpha,\gamma} \leq 1\}}(z),$$

the convex conjugate of $F(y) = F(z, r) = \frac{\lambda}{2}\|r - d\|_2^2 + \|z\|_{1,\gamma,\beta}$.

The last line in the reformulation of Eq. 2 as above now defines a saddle-point problem in a form that can be solved with the primal-dual algorithm as described in (40, 45, 57). The resulting iterations are provided in Algorithm 1. Note that there, instead of using fixed step sizes $\sigma$ and $\tau$, we employ an adaptive step-size choice as described in (46). The adaptive choice still ensures convergence but potentially allows larger step-sizes and hence a faster method. This is realized by the mapping $\mathcal{S}$, which is for $\theta \in (0, 1)$ defined as

$$\mathcal{S}(\sigma\tau, \mathfrak{n}) = \begin{cases} \mathfrak{n} & \text{if } \sqrt{\theta\sigma\tau} \geq \mathfrak{n}, \\ \sqrt{\theta\sigma\tau} & \text{if } \sqrt{\sigma\tau} \geq \mathfrak{n} > \sqrt{\theta\sigma\tau}, \\ \sqrt{\sigma\tau} & \text{else.} \end{cases} \quad [7]$$

The operators $P_\eta$, for $\eta > 0$, and $P_{L^2}$ in the algorithm correspond to the proximal mapping of $F^*$ and are given by

$$P_\eta(\xi)_{i,j,f} = \frac{\xi_{i,j,f}}{\max\left(1, \frac{|\xi_{i,j,f}|}{\eta}\right)} \quad \text{and} \quad P_{L^2}(\xi) = \frac{\xi - \sigma d}{1 + \frac{\sigma}{\lambda}},$$

where, abusing notation, $|\nu|$ is defined as $|\nu| = \sqrt{|\nu^1|^2 + |\nu^2|^2 + |\nu^3|^2}$ for $\nu \in \mathbb{R}^3$ and as $|\nu| = \sqrt{|\nu^1|^2 + |\nu^2|^2 + |\nu^3|^2 + 2|\nu^4|^2 + 2|\nu^5|^2 + 2|\nu^6|^2}$ for $\nu \in \mathbb{R}^6$. The divergence operators $\mathrm{div}^1_{\beta_i}$ and $\mathrm{div}^2_{\beta_i}$ are defined as the negative adjoints of $\nabla_{\beta_i}$ and $\mathcal{E}_{\beta_i}$, respectively.

**Initialize**: $(u, w_1, v, w_2)$, $(\bar{u}, \bar{w}_1, \bar{v}, \bar{w}_2)$, $(p_1, q_1, p_2, q_2, r)$, $\sigma, \tau > 0$

**Iterate**:

Dual Update:

$p_1 \leftarrow P_{\gamma_1\alpha_1}\left(p_1 + \sigma(\nabla_{\beta_1}(\bar{u} - \bar{v}) - \bar{w}_1)\right)$

$q_1 \leftarrow P_{\gamma_1\alpha_0}\left(q_1 + \sigma\mathcal{E}_{\beta_1}\bar{w}_1\right)$

$p_2 \leftarrow P_{\gamma_2\alpha_1}\left(p_2 + \sigma(\nabla_{\beta_2}\bar{v} - \bar{w}_2)\right)$

$q_2 \leftarrow P_{\gamma_2\alpha_0}\left(q_2 + \sigma\mathcal{E}_{\beta_2}\bar{w}_2\right)$

$r \leftarrow P_{L^2}\left(r + \sigma K\bar{u}\right)$

Primal Update:

$u^+ \leftarrow u - \tau\left(-\mathrm{div}^1_{\beta_1} p_1 + K^* r\right)$

$w_1^+ \leftarrow w_1 - \tau\left(-p_1 - \mathrm{div}^2_{\beta_1} q_1\right)$

$v^+ \leftarrow v - \tau\left(-\mathrm{div}^1_{\beta_1} p_1 - \mathrm{div}^1_{\beta_2} p_2\right)$

$w_2^+ \leftarrow w_2 - \tau\left(-p_2 - \mathrm{div}^2_{\beta_2} q_2\right)$

Step size Update:

$\sigma_+ \leftarrow \mathcal{S}\left(\sigma\tau, \frac{\|(u^+, w_1^+, v^+, w_2^+) - (u, w_1, v, w_2)\|_2}{\|H((u^+, w_1^+, v^+, w_2^+) - (u, w_1, v, w_2))\|_2}\right)$

$\tau_+ \leftarrow \sigma_+$

Extrapolation and Update:

$(\bar{u}, \bar{w}_1, \bar{v}, \bar{w}_2) \leftarrow 2(u^+, w_1^+, v^+, w_2^+) - (u, w_1, v, w_2)$

$(u, w_1, v, w_2) \leftarrow (u^+, w_1^+, v^+, w_2^+)$

**Algorithm 1:** Primal-dual algorithm for solving ICTGV regularized dynamic MR reconstruction.

1. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: Sensitivity encoding for fast MRI. Magn Reson Med 1999;42(5):952–962.
2. Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Reson Med 2002;47(6):1202–1210.
3. Stollberger R, Fazekas F. Improved perfusion and tracer kinetic imaging using parallel imaging. Topics in Magnetic Resonance Imaging 2004;15(4):245–254.
4. Kellman P, Epstein FH, McVeigh ER. Adaptive sensitivity encoding incorporating temporal filtering (TSENSE). Magn Reson Med 2001;45(5):846–852.
5. Tsao J, Boesiger P, Pruessmann KP. k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations. Magn Reson Med 2003;5:1031–1042.
6. Breuer FA, Kellman P, Griswold MA, Jakob PM. Dynamic autocalibrated parallel imaging using temporal GRAPPA (TGRAPPA). Magn Reson Med 2005;53(4):981–985.
7. Huang F, Akao J, Vijayakumar S, Duensing GR, Limkeman M. k-t GRAPPA: A k-space implementation for dynamic MRI with high reduction factor. Magn Reson Med 2005;54(5):1172–1184.
8. Candès EJ, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory 2006;52(2):489–509.
9. Donoho DL. Compressed sensing. IEEE Transactions on Information Theory 2006;52(4):1289–1306.
10. Jung H, Sung K, Nayak KS, Kim EY, Ye JC. k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI. Magn Reson Med 2009;61(1):103–116.
11. Otazo R, Kim D, Axel L, Sodickson DK. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. Magn Reson Med 2010;64(3):767–776.
12. Liu J, Lefebvre A, Zenge MO, Schmidt M, Mueller E, Nadar MS. 2D bSSFP real-time cardiac CINE-MRI: Compressed sensing featuring weighted redundant Haar wavelet regularization in space and time. Journal of Cardiovascular Magnetic Resonance 2013;15(1):1–2.
13. Adluru G, Awate SP, Tasdizen T, Whitaker RT, DiBella EV. Temporally constrained reconstruction of dynamic cardiac perfusion MRI. Magn Reson Med 2007;57(6):1027–1036.
14. Lustig M, Santos JM, Donoho DL, Pauly JM. k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity. In Proceedings of the 14th Annual Meeting of ISMRM, Seattle. 2006; 2420.
15. Feng L, Srichai MB, Lim RP, Harrison A, King W, Adluru G, Dibella EVR, Sodickson DK, Otazo R, Kim D. Highly accelerated real-time cardiac cine MRI using k-t SPARSE-SENSE. Magn Reson Med 2013;70(1):64–74.
16. Liang ZP. Spatiotemporal imaging with partially separable functions. In Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Washington, DC, USA, 2007; 988–991.
17. Brinegar C, Wu YJL, Foley LM, Hitchens TK, Ye Q, Ho C, Liang ZP. Real-time cardiac MRI without triggering, gating, or breath holding. In Engineering in Medicine and Biology Society, 2008. 30th Annual

International Conference of the IEEE. 2008; 3381–3384.

18. Pedersen H, Kozerke S, Ringgaard S, Nehrke K, Kim WY. k-t PCA: Temporally constrained k-t BLAST reconstruction using principal component analysis. Magn Reson Med 2009;62(3):706–716.
19. Velikina JV, Samsonov AA. Reconstruction of dynamic image series from undersampled MRI data using data-driven model consistency condition (MOCCO). Magn Reson Med 2015;74(5):1279–1290.
20. Candès EJ, Recht B. Exact matrix completion via convex optimization. Foundations of Computational Mathematics 2009;9(6):717–772.
21. Haldar JP, Liang ZP. Spatiotemporal imaging with partially separable functions: A matrix recovery approach. In IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Rotterdam, 2010; 716–719.
22. Lingala SG, Hu Y, Dibella E, Jacob M. Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR. Transactions on Medical Imaging 2011;30(5):1042–1054.
23. Zhao B, Haldar JP, Christodoulou AG, Liang ZP. Image reconstruction from highly undersampled. (k,t)-space data with joint partial separability and sparsity constraints. IEEE Transactions on Medical Imaging 2012;31(9):1809–1820.
24. Christodoulou AG, Zhang H, Zhao B, Hitchens T, Ho C, Liang ZP. High-resolution cardiovascular MRI by integrating parallel imaging with low-rank and sparse modeling. IEEE Transactions on Biomedical Engineering 2013;60(11):3083–3092.
25. Candès EJ, Li X, Ma Y, Wright J. Robust Principal Component Analysis? Journal of the ACM 2011; 58(3):1–37.
26. Ji H, Huang S, Shen Z, Xu Y. Robust Video Restoration by Joint Sparse and Low Rank Matrix Approximation. SIAM Journal on Imaging Sciences 2011;4(4):1122–1142.
27. Gao H, Rapacchi S, Wang D, Moriarty J, Meehan C, Sayre J, Laub G, Finn P, Hu P. Compressed sensing using prior rank, intensity and sparsity model (PRISM): Applications in cardiac cine MRI. In Proceedings of the 20th Annual Meeting of ISMRM, Melbourne. 2012; 2242.
28. Otazo R, Candès E, Sodickson DK. Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. Magn Reson Med 2015;73(3):1125–1136.
29. Tremoulheac B, Dikaios N, Atkinson D, Arridge SR. Dynamic MR Image Reconstruction Separation From Undersampled k,t-Space via Low-Rank Plus Sparse Prior. IEEE Transactions on Medical Imaging 2014;33(8):1689–1701.
30. Akçakaya M, Basha TA, Pflugi S, Foppa M, Kissinger KV, Hauser TH, Nezafat R. Localized spatio-temporal constraints for accelerated CMR perfusion. Magn Reson Med 2014;72(3):629–639.
31. Ong F, Zhang T, Cheng J, Uecker M, Lustig M. Beyond low rank + sparse: Multi-scale low rank reconstruction for dynamic contrast enhanced imaging. In Proceedings of the 23th Annual Meeting of ISMRM, Toronto. 2015; 0575.
32. Yoon H, Kim SK, Kim D, Bresler Y, Ye JC. Motion Adaptive Patch-Based Low-Rank Approach for

Comressed Sensing Cardiac Cine MRI. IEEE Transactions on Medical Imaging 2014;33(11):2069–2085.

33. Chen X, Salerno M, Yang Y, Epstein FH. Motion-Compensated Compressed Sensing for Dyamic Contrast-Enhanced MRI Using Regional Spatiotemporal Sparsity and Region Tracking: Block LO-rank Sparsity with Motion-guidance (BLOSM). Magn Reson Med 2014;72(4):1028–1038.
34. Rizwan A, Xue H, Giri S, Ding Y, Craft J, Simonetti OP. Variable density incoherent spatiotemporal acquisition (VISTA) for highly accelerated cardiac MRI. Magn Reson Med 2015;74(5):1266–1278.
35. Tsai CM, Nishimura DG. Reduced aliasing artifacts using variable-density k-space sampling trajectories. Magn Reson Med 2000;43:452–458.
36. Winkelmann S, Schaeffter T, Koehler T, Eggers H, Doessel O. An optimal radial profile order based on the golden ratio for time-resolved MRI. IEEE Transactions on Medical Imaging 2007;26(1):68–76.
37. Bredies K, Kunisch K, Pock T. Total generalized variation. SIAM Journal on Imaging Sciences 2010; 3(3):492–526.
38. Bredies K, Holler M. Regularization of linear inverse problems with total generalized variation. Journal of Inverse and Ill-Posed Problems 2014;22(6):871–913.
39. Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. Journal of Physics D 1992;60:259–268.
40. Knoll F, Bredies K, Pock T, Stollberger R. Second order total generalized variation (TGV) for MRI. Magn Reson Med 2011;65(2):480–491.
41. Holler M, Kunisch K. On Infimal Convolution of TV Type Functionals and Applications to Video and Image Reconstruction. SIAM Journal on Imaging Sciences 2014;7(4):2258–2300.
42. 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(6):1086–1098.
43. Fessler J, Sutton B. Nonuniform fast fourier transforms using min-max interpolation. IEEE Transactions on Signal Processing 2003;51(2):560–574.
44. Schloegl M, Holler M, Bredies K, Stollberger R. A variational approach for coil-sensitivity estimation for undersampled phase-sensitive dynamic MRI reconstruction. In Proceedings of the 23th Annual Meeting of ISMRM, Toronto. 2015; 3692.
45. Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathemtical Imaging and Vision 2011;40(1):120–145.
46. Bredies K, Holler M. TGV-based framework for variational image decompression, zooming and reconstruction. Part II:Numerics. SIAM Journal in Imaging Sciences 2015;8(4):2851–2886.
47. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image qualifty assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing 2004;13(4):600–612.
48. Freiberger M, Knoll F, Bredies K, Scharfetter H, Stollberger R. The AGILE library for image reconstruction in biomedical sciences using graphics card hardware acceleration. Computing in Science and Engineering 2013;15:34–44.

49. Wissman L, Santelli C, Segars WP, Kozerke S. MRXCAT: Realistic numerical phantoms for cardiovascular magnetic resonance. Journal of Cardiovascular Magnetic Resonance 2014;16(1).
50. Velikina JV, Alexander AL, Samsonov AA. Accelerating MR parameter mapping using sparsity-promoting regularization in parametric dimension. Magn Reson Med 2013;70(5):1263–1273.
51. Zhao B, Lu W, Hitchens TK, Lam F, Ho C, Liang ZP. Accelerated MR parameter mapping with low-rank and sparsity constraints. Magn Reson Med 2015;74(2):489–498.
52. Inati SJ, Naegele JD, Zwart NR, Roopchansingh V, Lizak MJ, Hansen DC, Liu CY, Atkinson D, Kellman P, Kozerke S, Xue H, Campbell-Washburn AE, Sørensen TS, Hansen MS. ISMRM Raw data format: A proposed standard for MRI raw datasets. Magn Reson Med 2016;.
53. Chandarana H, Block TK, Rosenkrantz AB, Lim RP, Kim D, Mossa DJ, Babb JS, Kiefer B, Lee VS. Free-breathing radial 3D fat-suppressed T1-weighted gradient echo sequence: A viable alternative for contrast-enhanced liver imaging in patients unable to suspend respiration. Investigative Radiology 2011; 46(10):648–653.
54. Piccini D, Littmann A, Nielles-Vallespin S, Zenge MO. Spiral phyllotaxis: The natural way to construct a 3D radial trajectory in MRI. Magn Reson Med 2011;66(4):1049–1056.
55. Bredies K, Sun H. Preconditioned Douglas-Rachford algorithms for TV and TGV regularized variational imaging problems. Journal of Mathematical Imaging and Vision 2015;52(3):317–344.
56. Bredies K, Sun H. Preconditioned Douglas-Rachford splitting methods for convex-concave saddle-point problems. SIAM Journal on Numerical Analysis 2015;53(1):421–444.
57. Bredies K. Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty. In A Bruhn, T Pock, XC Tai, eds., Efficient Algorithms for Global Optimization Methods in ComputerVision, volume 8293 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2014; 44–77.

## LIST OF FIGURES

- Figure 1: Evaluation of the model parameters $s, t_1, t_2$ by means of RMSE (solid lines) and SSIM (dashed lines) for a CINE cardiac test-case with acceleration factors $r = 5$ (a) and $r = 10$ (b). The colors indicate three different choices for s. The horizontal axis show different values for $t_1$ and the vertical axis the corresponding best RMSE and SSIM values achieved with $t_2 \in \{0.2, 0.4, 0.5, 1, 2.5, 3\}$, which are marked (arrow) with the corresponding best value for $t_2$.

- Figure 2: (a) RMSE (blue curves) and SSIM (red curves) evaluation exemplified for one CINE cardiac test case for different regularization parameters $\lambda$ and acceleration factors $r = 4$ (solid line), $r = 8$ (dashed line), $r = 12$ (dashed-dotted line) and $r = 16$ (dotted line). The corresponding optimal values for $\lambda$ are indicated with black circles and were calculated by spline-interpolation between the used sample-points. (b) Optimal values for $\lambda$ according to RMSE (blue) and SSIM (red) for different acceleration factors and the test case displayed in (a) (squares) and the second test case (dots). The linear regression to both cases and metrics (black dashed line) yields the proposed parameter choice.

- Figure 3: Magnitude images from simulated accelerations $r = (4, 8, 12, 16)$ for the four-chamber-view bSSFP dataset. The fully sampled sum-of-squares reconstruction is displayed in the 1st column and the therein indicated time-lines are shown in the 2nd and 3rd row.

- Figure 4: Comparison of magnitude images of fully sampled reference reconstruction (1st column), L+S (2nd and 3rd column) and ICTGV reconstruction (4th and 5th column) for bSSFP CINE cardiac data acquired in short-axis-view and undersampling factors of $r = (12, 16)$. A late diastolic time-frame is displayed in the first row and indicated vertical and horizontal time-lines in the 2nd and 3rd row. A closeup of the heart region is displayed in the 4th row.

- Figure 5: Comparison of magnitude images of fully sampled reference reconstruction (1st column), L+S (2nd and 3rd column) and ICTGV reconstruction (4th and 5th column) for the cardiac perfusion dataset and undersampling factors of $r = (12, 16)$. A selected time-frame is displayed in the first row and the indicated time-line in the second row. A close-up of the heart region is displayed in the third row.

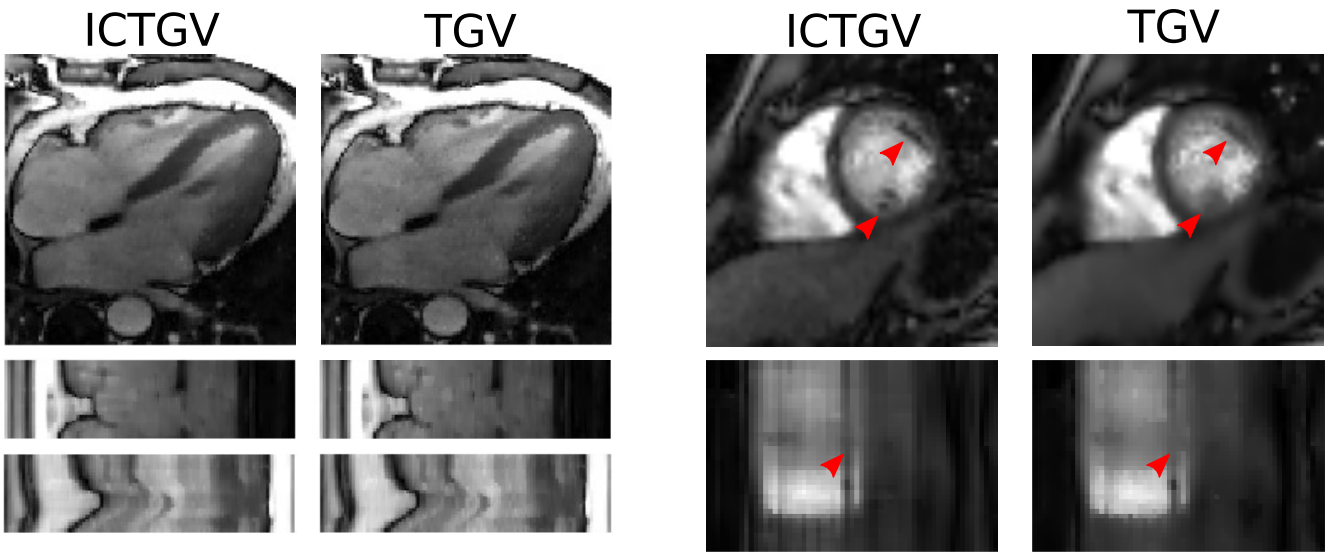
- Figure 6: Coil-combined fully sampled reference (1st column), ICTGV (2nd column), kt-RPCA (3rd column), kt-SLR (fourth column) and kt-FOCUSS (fifth column) single-coil reconstructions from undersampled Cartesian data ($r = 7.3$) for the cardiac perfusion dataset with computed SER values in dB. The mean time-course of an indicated $3 \times 3$ voxel region within the right ventricle is plotted for all methods under investigation with a closeup view of the peak-signal.

- Figure 7: Coil-combined fully sampled reference (1st column), ICTGV (2nd column), kt-RPCA (3rd column), kt-SLR (fourth column) and kt-FOCUSS (fifth column) single-coil reconstructions from

undersampled Cartesian data with $r = 3.6$ (1st to 3rd row) and $r = 7.5$ (4th to 6th row) for the four chamber CINE cardiac dataset. Reconstructions are displayed with a selected time-frame and the indicated horizontal and vertical time-lines.

## SUPPORTING INFORMATION

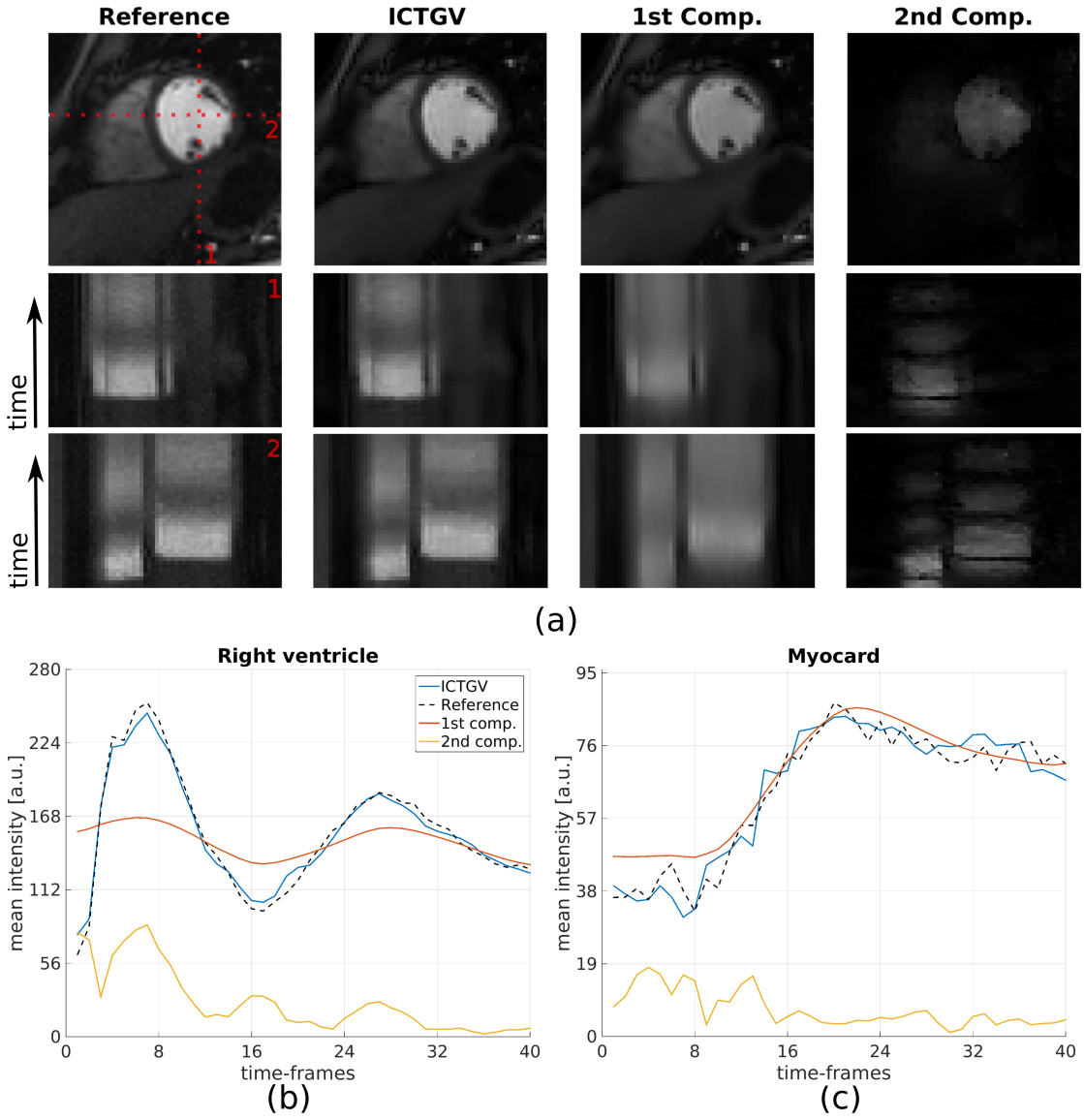
- Supporting Figure S1: Comparison of ICTGV vs TGV multi-coil reconstruction for CINE cardiac imaging (1st and 2nd column, $r = 16$) and cardiac perfusion imaging (3rd and 4th column, $r = 12$).
- Supporting Figure S2: (a) Fully sampled reference (1st column), ICTGV reconstruction (2nd column), 1st component (3rd column) and 2nd component (4th column) for a fixed reduction factor of $r = 8$ and a selected time-frame of a short axis perfusion dataset with the corresponding horizontal and vertical time-lines in the 2nd and 3rd row as indicated in the reference frame. Mean intensity change (magnitude) over time within the right ventricle (b) and the myocardium (c) due to the contrast agent for the reference (black-dotted line), ICTGV reconstruction (blue solid line) and components (red and yellow solid line). The second component was rescaled for display purposes.
- Supporting Figure S3: Quantitative evaluation of reconstruction results for different choices of the time-space-weighting $t$, for spatio-temporal TV (first and second column) and spatio-temporal TGV (third and fourth column), by means of RMSE and SSIM.
- Supporting Table S1: Quantitative evaluation (best values underlined) of single-coil reconstruction for ICTGV against kt-RPCA and kt-SLR and kt-FOCUSS reconstruction for CINE cardiac and cardiac perfusion cases by means of SSIM and SER against the fully sampled single-coil reconstruction.



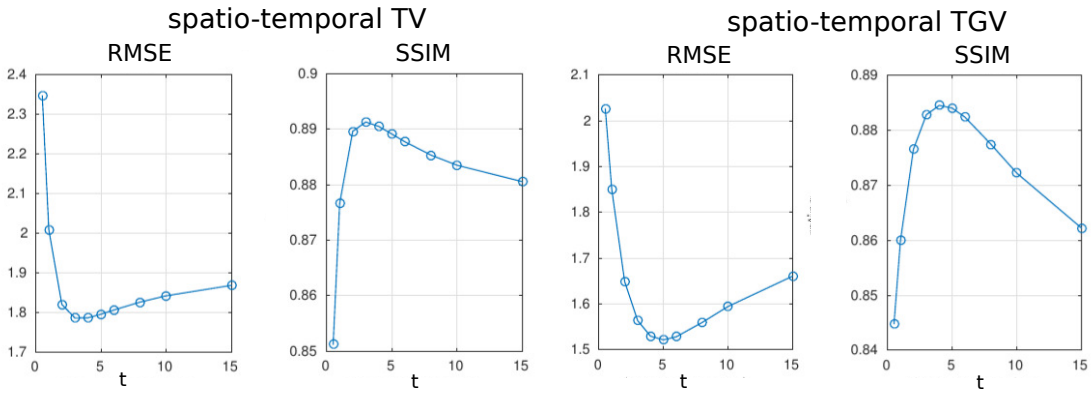

Supplementary Figure 1: Comparison of ICTGV vs TGV multi-coil reconstruction for CINE cardiac imaging (1st and 2nd column, $r = 16$) and cardiac perfusion imaging (3rd and 4th column, $r = 12$).

Supplementary Table 1: Quantitative evaluation (best values underlined) of single-coil reconstruction for ICTGV against kt-RPCA and kt-SLR and kt-FOCUSS reconstruction for CINE cardiac and cardiac perfusion cases by means of SSIM and SER against the fully sampled single-coil reconstruction.

| | ICTGV | | kt-RPCA | | kt-SLR | | kt-FOCUSS | |
|---|---|---|---|---|---|---|---|---|
| | SER (dB) | SSIM | SER (dB) | SSIM | SER (dB) | SSIM | SER (dB) | SSIM |
| Four Chamber View | | | | | | | | |
| $r = 1.96$ | <u>29.24</u> | <u>0.9217</u> | 27.71 | 0.9214 | 28.86 | 0.9169 | 27.49 | 0.9052 |
| $r = 3.6$ | <u>25.66</u> | <u>0.8524</u> | 21.63 | 0.8216 | 24.85 | 0.8295 | 24.56 | 0.8436 |
| $r = 5.2$ | <u>23.42</u> | <u>0.8047</u> | 19.95 | 0.7674 | 22.52 | 0.7759 | 22.58 | 0.7955 |
| $r = 6.5$ | <u>22.30</u> | <u>0.7680</u> | 18.70 | 0.7260 | 21.36 | 0.7335 | 20.87 | 0.7429 |
| $r = 7.5$ | <u>20.98</u> | <u>0.7274</u> | 17.87 | 0.6879 | 20.29 | 0.6920 | 19.83 | 0.7058 |
| Cardiac Perfusion | | | | | | | | |
| $r = 1.9$ | <u>27.19</u> | <u>0.9429</u> | 25.99 | 0.9394 | 26.90 | 0.9421 | 26.63 | 0.9420 |
| $r = 3.5$ | <u>24.73</u> | <u>0.9096</u> | 23.14 | 0.8980 | 23.94 | 0.8979 | 23.37 | 0.8946 |
| $r = 4.8$ | <u>23.23</u> | <u>0.8847</u> | 21.41 | 0.8662 | 21.95 | 0.8562 | 21.59 | 0.8614 |
| $r = 5.7$ | <u>22.16</u> | <u>0.8713</u> | 20.45 | 0.8505 | 20.76 | 0.8302 | 20.54 | 0.8415 |
| $r = 6.5$ | <u>21.64</u> | <u>0.8614</u> | 19.93 | 0.8391 | 19.40 | 0.7780 | 20.14 | 0.8325 |
| $r = 7.3$ | <u>20.62</u> | <u>0.8402</u> | 18.69 | 0.8118 | 18.50 | 0.7500 | 18.83 | 0.7988 |




Supplementary Figure 2: (a) Fully sampled reference (1st column), ICTGV reconstruction (2nd column), 1st component (3rd column) and 2nd component (4th column) for a fixed reduction factor of $r = 8$ and a selected time-frame of a short axis perfusion dataset with the corresponding horizontal and vertical time-lines in the 2nd and 3rd row as indicated in the reference frame. Mean intensity change (magnitude) over time within the right ventricle (b) and the myocardium (c) due to the contrast agent for the reference (black-dotted line), ICTGV reconstruction (blue solid line) and components (red and yellow solid line). The second component was rescaled for display purposes.




Supplementary Figure 3: Quantitative evaluation of reconstruction results for different choices of the time-space-weighting $t$, for spatio-temporal TV (first and second column) and spatio-temporal TGV (third and fourth column), by means of RMSE and SSIM.