Accelerated Dynamic MRI Exploiting Sparsity and Low-Rank Structure: k-t SLR

Sajan Goud Lingala*, *Student Member, IEEE*, Yue Hu, *Student Member, IEEE*, Edward DiBella, *Member, IEEE*, and Mathews Jacob, *Member, IEEE*

Method to **reconstruct dynamic MRI(cine MRI) data**

1) Classic method

Based on the sparsity or banded structure in Fourier space.

2) Proposed method

Based on the compact representation in **Karhunen Louve transform(KLT) domain.**

And it uses spectrally regularized reconstruction.

1.
Introduction

1. Introduction

Achieving high spatiotemporal resolutions is challenging in dynamic MRI.

Preceding models, which exploit the banded structure or sparsity of the data in x-f space, have noted poor performance under respiratory motion.

Several researchers have suggested Karhunen Louve transform(**KLT**) as an alternative to x-f space.

> ***KLT*** *: A data-derived transform. The resulting adaptive scheme is capable of exploiting the correlations in the data, even when the temporal profiles of the voxels are not periodic.*

Current KLT-based algorithms rely on a two-step approach.

1) Estimate the **temporal basis** function using the SVD of a training dataset.

The training dataset is an image time series with low spatial resolution and Nyquist temporal sampling rate.

The training dataset is obtained as the IFFT of the central phase encodes, which is collected along with dense k-space samples at sub-Nyquist temporal sampling rates.

2) Reconstruct the data from sub-Nyquist sampled k-space data using estimated temporal basis.

Proposal

Simultaneous estimation of the temporal basis functions and its spatial weights directly from the entire k-t space data.

It is enabled by the reinterpretation of the KLT based reconstruction as a **spectrally regularized matrix recovery scheme**.

Recover a **low-rank matrix** using **nuclear norm**.

Nonconvex spectral penalties to minimize the number of measurements required to covert a low-rank matrix.

Additionally exploit the **sparsity** of the matrix to improve the *recovery rate*.

Variable splitting algorithm for the fast minimization of the optimization criterion.

2.
Background

Notations

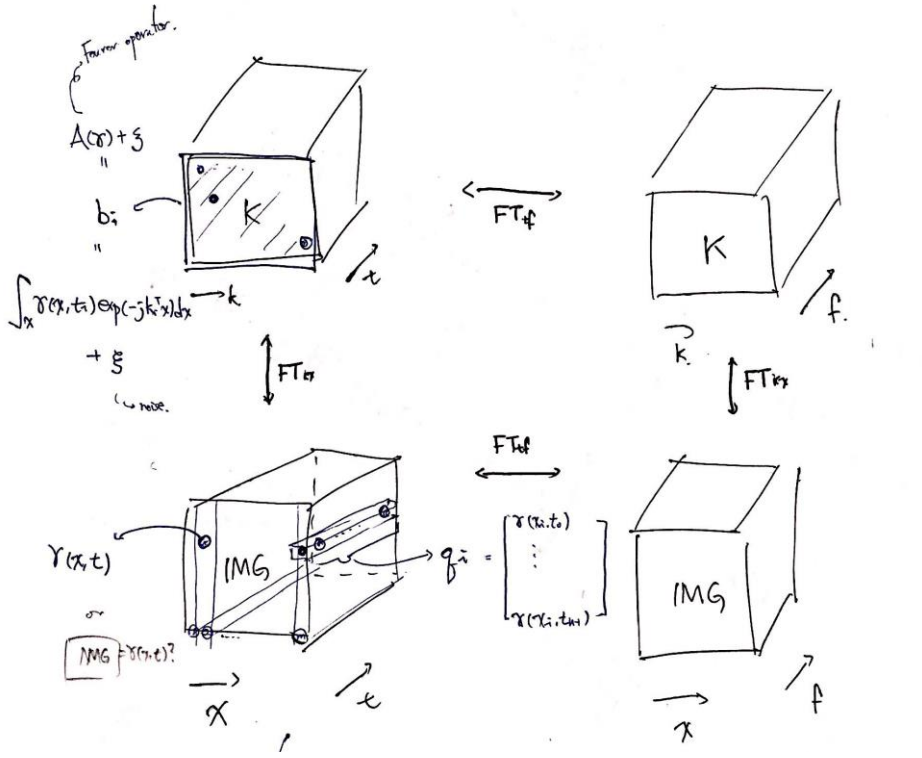
Spatiotemporal signal: $\gamma(\mathbf{x}, t)$

Temporal profiles of the voxels: $\mathbf{q}_i$

$$= [\gamma(\mathbf{x}_i, t_0), \gamma(\mathbf{x}_i, t_1), \ldots, \gamma(\mathbf{x}_i, t_{n-1})]^T; \quad i = 0, \ldots, m-1$$

Spatiotemporal signal in a matrix form:

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma(\mathbf{x}_0, t_0) & \ldots & \gamma(\mathbf{x}_0, t_{n-1}) \\ \vdots & & \\ \gamma(\mathbf{x}_{m-1}, t_0) & \ldots & \gamma(\mathbf{x}_{m-1}, t_{n-1}) \end{bmatrix}$$



Measurements: $\mathbf{b}_i$

$$= \int_{\mathbf{x}} \gamma(\mathbf{x}, t_i) \exp\left(-j\mathbf{k}_i^T\mathbf{x}\right)\, d\mathbf{x} + \mathbf{n}_i; \quad i = 0, \ldots, s-1.$$

*The set of sampling **locations**:*

$$\Xi = \{(\mathbf{k}_i, t_i), i = 0, \ldots, s-1\}$$

Vector form: $\mathbf{b} = \mathcal{A}(\gamma) + \mathbf{n}$

where, $\mathcal{A}$ is the Fourier sampling operator

Goal:

To recover the signal $\gamma(\boldsymbol{x}, t)$ from the measured k-t space samples.

2.1. Dynamic MRI Using KLT

Goal: To recover the signal $\gamma(\boldsymbol{x}, t)$ from the measured k-t space samples.

Spatiotemporal signal matrix(Γ, Casorati matrix) can be decomposed using SVD.

$$\mathbf{\Gamma} = \underbrace{\mathbf{U}}_{m\times r}\underbrace{\mathbf{\Sigma}}_{r\times r}\underbrace{\mathbf{V}^H}_{r\times n}.$$

Since the rows of are linearly dependent, the rank, $r < \min(m, n)$

The spatiotemporal signal(γ) can be expressed as a spatially weighted(ρ_i) linear combination of temporal basis functions(v_i).

$$\gamma(\mathbf{x}, t) = \sum_{i=0}^{r-1} \rho_i(\mathbf{x})\, v_i(t)$$

$$\Gamma = US \cdot V^H = \begin{bmatrix} \rho_{i=1}(x_{=1}) & \cdots & \rho_{i=r}(x_{=1}) \\ \vdots & \ddots & \vdots \\ \rho_{i=1}(x_{=m}) & \cdots & \rho_{i=r}(x_{=m}) \end{bmatrix} \begin{bmatrix} v_{i=1}(t_{=1}) & \cdots & v_{i=1}(t_{=n}) \\ \vdots & \ddots & \vdots \\ v_{i=r}(t_{=1}) & \cdots & v_{i=r}(t_{=n}) \end{bmatrix}$$

2.1. Dynamic MRI Using KLT

Most of the KLT-based algorithms strategy

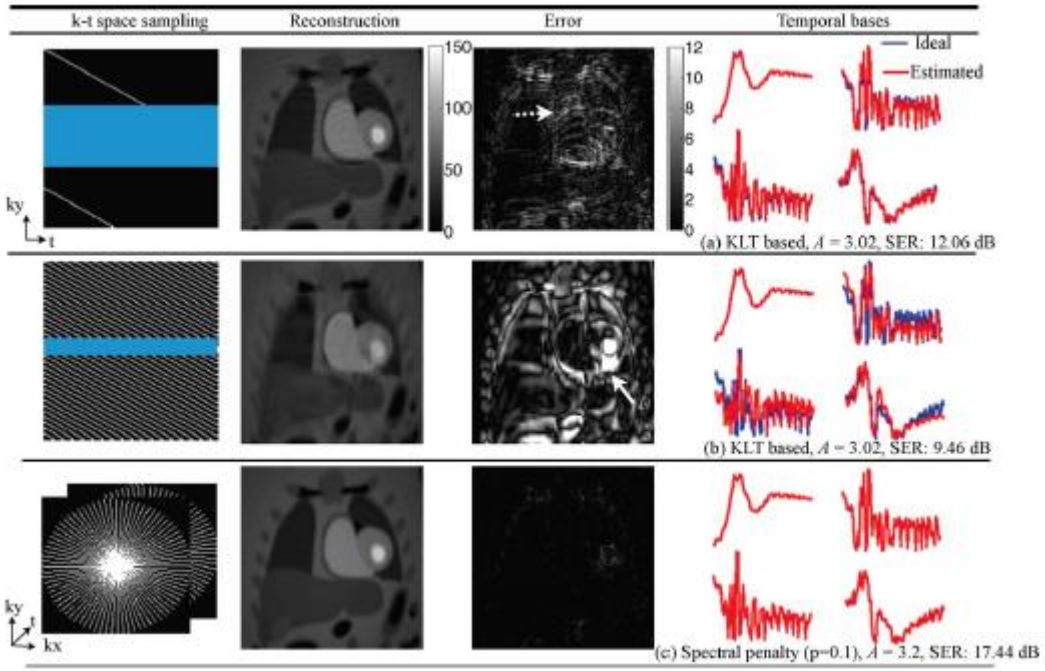
1) Estimate the temporal basis function(v_i) using SVD of the training image time-series.

The training data consists of dynamic image data, acquired with low spatial resolution.
(few central phase encoding lines with Nyquist temporal sampling rate)

2) Recover the cardiac data from sub-Nyquist measurements using estimated temporal basis functions.

Since $r \ll n$, this approach provides a significant reduction in the number of unknowns.

This scheme is violated when the number of phase encodes in the training data are too few.



While the acquisition of more training data can minimize these problems, this comes at the expense of the number of high-frequency encodes that can be acquired at a specified acceleration rate; this can often result in aliasing artifacts.

Fig. 6. Comparison of the two-step KLT schemes (two top rows) with the spectrally regularized reconstruction scheme ($p = 0.1, \lambda_2 = 0$),

In summary, the performance of the two-step strategy requires a fine balance between the amount of training data and the number of high-frequency encodes.

2.2. Matrix Recovery Using Nuclear Norm Minimization

A low rank matrix can be recovered by solving below **constrained optimization** problem.

$$\mathbf{\Gamma}^* = \arg\min_{\mathbf{\Gamma}} ||\mathcal{A}(\mathbf{\Gamma}) - \mathbf{b}||^2 \text{ such that } \text{rank}\,(\mathbf{\Gamma}) \leq r.$$

Reformulating into **Lagrange's multiplier** framework,

$$\mathbf{\Gamma}^* = \arg\min_{\mathbf{\Gamma}} ||\mathcal{A}(\mathbf{\Gamma}) - \mathbf{b}||^2 + \lambda\, \text{rank}\,(\mathbf{\Gamma}).$$

Since the rank penalty is nonconvex, it is often replace with the **nuclear norm**, which is the closest convex relaxation.

$$\mathbf{\Gamma}^* = \arg\min_{\mathbf{\Gamma}} \underbrace{||\mathcal{A}(\mathbf{\Gamma}) - \mathbf{b}||^2 + \lambda\, ||\mathbf{\Gamma}||_*}_{\mathcal{C}(\mathbf{\Gamma})}.$$

nuclear norm of r-rank matrix
A = USV*, denoted by
$||A||_* = \sum_i S_{i,i}$

3.
k-t SLR Formulation

3.1. Regularized Matrix Recovery Using Spectral Priors

A spectrally regularized optimization problem

$$\mathbf{\Gamma}^* = \arg\min_{\mathbf{\Gamma}} \|\mathcal{A}(\mathbf{\Gamma}) - \mathbf{b}\|^2 + \lambda\,\varphi(\mathbf{\Gamma})$$

where
$\varphi(\Gamma)$: spectral penalty

$$\varphi(\Gamma) = \left(\left|\left|\Gamma\right|\right|_p\right)^p = \sum_{i=1}^{\min\{m,n\}} \sigma_i^p$$

It means the nuclear norm for p=1

Note:

The cost function does not depend explicitly on the temporal basis functions or its spatial weights as in the case of 2-step KLT schemes.

However, the it implicitly iteratively updates the temporal basis functions and spatial weights.

In addition, if the intensity profiles of the voxels are periodic, the columns of **Γ** may be sparse in the Fourier domain.

3.2. Regularized Matrix Recovery Using Spectral and Sparsity Priors

We **exploit the sparsity of the signal** in specified basis sets to further improve the recovery rate.

Consider the matrix $\Gamma \in R^{mXn}$ that has at most N nonzero entries in a specified basis:

$$\left|\left|\boldsymbol{\Phi^H \Gamma \Psi}\right|\right|_{\boldsymbol{l_0}} \leq \boldsymbol{N}.$$

where
Φ : 2−D wavelet transform to sparsify row−space (each of the images in the time series) of Γ.

Ψ : 1−D Fourier transform to sparsify column−space (exploit the pseudo−periodic nature of motion) of Γ.

$$\Phi^H \Gamma \Psi = \Phi^H US \cdot V^H \Psi$$

$$U: column_{+left\ null}\ space's\ eigenvectors\ \ of\ \boldsymbol{\Gamma}$$
$$V: row_{+null}\ space's\ eigenvectors\ of\ \boldsymbol{\Gamma}$$

$$\gamma(\mathbf{x}, t) = \sum_{i=0}^{r-1} \rho_i(\mathbf{x})\, v_i(t)$$

$$\Gamma = US \cdot V^H = \begin{bmatrix} \rho_{i=1}(x_{=1}) & \cdots & \rho_{i=r}(x_{=1}) \\ \vdots & \ddots & \vdots \\ \rho_{i=1}(x_{=m}) & \cdots & \rho_{i=r}(x_{=m}) \end{bmatrix} \begin{bmatrix} v_{i=1}(t_{=1}) & \cdots & v_{i=1}(t_{=n}) \\ \vdots & \ddots & \vdots \\ v_{i=r}(t_{=1}) & \cdots & v_{i=r}(t_{=n}) \end{bmatrix}$$

The use of this prior knowledge can significantly reduce the number of DOF and the number of measurements required to recover the matrix.

3.2. Regularized Matrix Recovery Using Spectral and Sparsity Priors

The cost function which exploits sparsity and low-rank properties

$$\mathbf{\Gamma}^* = \arg\min_{\mathbf{\Gamma}} \|\mathcal{A}(\mathbf{\Gamma}) - \mathbf{b}\|^2$$
$$\text{s.t } \{\text{rank}(\mathbf{\Gamma}) \leq \text{r}, \|\mathbf{\Phi}^H \mathbf{\Gamma} \mathbf{\Psi}\|_{\ell_0} < K\}$$

Using Lagrange's multipliers and relaxing the penalties,

$$\mathbf{\Gamma}^* = \arg\min_{\mathbf{\Gamma}} \|\mathcal{A}(\mathbf{\Gamma}) - \mathbf{b}\|^2 + \lambda_1 \varphi(\mathbf{\Gamma}) + \lambda_2 \psi(\mathbf{\Gamma})$$

where
$\psi(\Gamma) = \left||\Phi^H \Gamma \Psi|\right|_{l_1}$ is a surrogate for the l0 term

$\varphi(\Gamma) = \left||\Gamma|\right|_p^p$ the cost function is convex and hence will have a unique minimum.

3.2. Regularized Matrix Recovery Using Spectral and Sparsity Priors

TV regularization adaptation

$$\psi(\mathbf{\Gamma}) = \left\| \sqrt{\sum_{i=0}^{q-1} \left| \mathbf{\Phi}_i^H \mathbf{\Gamma} \mathbf{\Psi}_i \right|^2} \right\|_{\ell_1}.$$

Where q=3

$$\Phi_0 = D_x \; and \; \Psi_0 = I$$

$$\Phi_1 = D_y \; and \; \Psi_1 = I$$

$$\Phi_2 = I \; and \; \Psi_2 = D_t$$

D : finite difference matrix

The proposed scheme is well posed since the sparsifying transforms are incoherent with the Fourier sampling operator.

We do not need the additional assumption of the right and the left singular vectors of **Γ** to be incoherent with the operator that picks the samples/matrix entries of **Γ** to make the problem well posed.

4.
Optimization Algorithm

4.1. Alternative Minimization

P1: Constrained minimization problem using variable splitting

$$\begin{aligned}\boldsymbol{\Gamma}^* = \arg\min_{\boldsymbol{\Gamma},\mathbf{R},\mathbf{S}} \|\mathcal{A}(\boldsymbol{\Gamma}) - \mathbf{b}\|^2 + \lambda_1 \varphi(\mathbf{R}) \\ + \lambda_2 \left\| \sqrt{\sum_{i=0}^{q-1} \|\mathbf{S}_i\|^2} \right\|_{\ell_1} \\ \text{s.t. } \boldsymbol{\Gamma} = \mathbf{R}; \mathbf{S}_i = \boldsymbol{\Phi}_i^H \boldsymbol{\Gamma} \boldsymbol{\Psi}_i; i = 0, \ldots, q-1.\end{aligned}$$

P2: Using the penalty method, where we minimize

$$\begin{aligned}\mathcal{D}_{\beta_1,\beta_2}(\boldsymbol{\Gamma}, \mathbf{R}, \mathbf{S}_i) = \|\mathcal{A}(\boldsymbol{\Gamma}) - \mathbf{b}\|^2 + \lambda_1 \varphi(\mathbf{R}) \\ + \lambda_2 \left\| \sqrt{\sum_{i=0}^{q-1} \|\mathbf{S}_i\|^2} \right\|_{\ell_1} \\ + \frac{\beta_1}{2} \|\boldsymbol{\Gamma} - \mathbf{R}\|^2 \\ + \frac{\beta_2}{2} \sum_{i=0}^{q-1} \|\boldsymbol{\Phi}_i^H \boldsymbol{\Gamma} \boldsymbol{\Psi}_i - \mathbf{S}_i\|^2\end{aligned}$$

P2: Using a three-step **alternating minimization** scheme

$$\begin{aligned}
\mathbf{\Gamma}_{n+1} &= \arg\min_{\mathbf{\Gamma}} \|\mathcal{A}(\mathbf{\Gamma}) - \mathbf{b}\|^2 + \frac{\beta_1}{2}\|\mathbf{\Gamma} - \mathbf{R}_n\|^2 \\
&\quad + \frac{\beta_2}{2}\sum_{i=0}^{q-1}\|\mathbf{\Phi}_i^H \mathbf{\Gamma}\mathbf{\Psi}_i - \mathbf{S}_{i,n}\|^2 \\
\mathbf{R}_{n+1} &= \arg\min_{\mathbf{R}} \|\mathbf{\Gamma}_{n+1} - \mathbf{R}\|^2 + 2\lambda_1/\beta_1 \varphi(\mathbf{R}) \\
\mathbf{S}_{i,n+1} &= \arg\min_{\{\mathbf{S}_i\}} \sum_{i=0}^{q-1} \left\|\mathbf{\Phi}_i^H \mathbf{\Gamma}_{n+1}\mathbf{\Psi} - \mathbf{S}_i\right\|^2 \\
&\quad + 2\lambda_2/\beta_2 \left\| \sqrt{\sum_{i=0}^{q-1}\|\mathbf{S}_i\|^2} \right\|_{\ell_1}; \\
&\qquad\qquad i = 0, \ldots, q-1.
\end{aligned}$$

4.1. Alternative Minimization

$$\mathbf{\Gamma}_{n+1} = \arg\min_{\mathbf{\Gamma}} \|\mathcal{A}(\mathbf{\Gamma}) - \mathbf{b}\|^2 + \frac{\beta_1}{2}\|\mathbf{\Gamma} - \mathbf{R}_n\|^2 + \frac{\beta_2}{2}\sum_{i=0}^{q-1}\|\mathbf{\Phi}_i^H \mathbf{\Gamma} \mathbf{\Psi}_i - \mathbf{S}_{i,n}\|^2$$

1st subproblem

- Quadratic form -> can get the analytic solution

$$\begin{aligned}\mathbf{\Gamma}_{n+1} &= \left(\mathcal{A}^T\mathcal{A} + \frac{\beta_1}{2}\mathcal{I} + \frac{\beta_2}{2}\sum_{i=0}^{q-1}\mathcal{Q}_i^T\mathcal{Q}_i\right)^{-1} \\ &\quad \times \left(\mathcal{A}^T\mathbf{b} + \frac{\beta_1}{2}\mathbf{R} + \frac{\beta_2}{2}\sum_{i=0}^{q-1}\mathbf{S}_i\right) \\ &= \mathcal{T}(\mathbf{R}, \mathbf{S}_i)\end{aligned}$$

where $Q_i(\Gamma) = \Phi_i^H \Gamma \Psi_i; _{i=0,\ldots,q-1}$

This step can be efficiently evaluated in the Fourier domain, if the measurements are sampled on a Cartesian grid. Otherwise, we uses a few CG gradient steps.

4.1. Alternative Minimization

$$\mathbf{R}_{n+1} = \arg\min_{\mathbf{R}} \|\mathbf{\Gamma}_{n+1} - \mathbf{R}\|^2 + 2\lambda_1/\beta_1 \varphi(\mathbf{R})$$

2nd subproblem

- Standard nuclear norm minimization problem
- Using a singular value thresholding of Γ_{n+1}

$$\mathbf{R}_{n+1} = (\mathcal{S}_{\lambda_1/\beta_1} \circ \mathcal{T})(\mathbf{R}_n, \mathbf{S}_{i,n})$$

where the singular value shrinkage is specified by

$$\mathcal{S}_{\lambda_1/\beta_1}(\mathbf{\Gamma}_{n+1}) = \sum_{i=0}^{\min(m,n)} \left(\sigma_i - \lambda\sigma_i^{p-1}/\beta\right)_+ \mathbf{u}_i\mathbf{v}_i^*.$$

$$(\sigma)_+ = \begin{cases} \sigma, & \text{if } \sigma \geq 0 \\ 0, & \text{else} \end{cases}.$$

Soft thresholding scheme if p=1

4.1. Alternative Minimization

$$\mathbf{S}_{i,n+1} = \arg\min_{\{\mathbf{S}_i\}} \sum_{i=0}^{q-1} \left\| \mathbf{\Phi}_i^H \mathbf{\Gamma}_{n+1} \mathbf{\Psi} - \mathbf{S}_i \right\|^2 + 2\lambda_2/\beta_2 \left\| \sqrt{\sum_{i=0}^{q-1} \|\mathbf{S}_i\|^2} \right\|_{\ell_1}; \quad i = 0, \ldots, q-1.$$

3rd subproblem

- It requires the joint processing of magnitude reduction on all the terms $Q_i(\Gamma_{n+1})_{;i=0,\ldots,q-1}$.
- This approach is termed as multidimensional shrinkage of $\{Q_i(\Gamma_{n+1})_{;i=0,\ldots,q-1}\}$.

$$\begin{aligned}\mathbf{S}_{i,n+1} &= \frac{\mathcal{Q}_i(\mathbf{\Gamma}_{n+1})}{\sum_{i=0}^{q-1} \|\mathcal{Q}_i(\mathbf{\Gamma}_{n+1})\|^2} \\ &\quad \cdot \left(\sum_{i=0}^{q-1} \|\mathcal{Q}_i(\mathbf{\Gamma}_{n+1})\|^2 - \frac{\lambda_2}{\beta_2} \right)_+ \\ &= \kappa_{\lambda_2/\beta_2}(\mathbf{\Gamma}_{n+1}).\end{aligned}$$

4.2. Convergence

The operator T and the shrinkage operations S, κ are non-expansive.

The convergence is guaranteed as $\beta_1, \beta_2 \rightarrow \infty$.

However, the quadratic problem of 1st subproblem will be become ill-conditioned for high β_1, β_2.

To overcome the tradeoff between computational complexity and accuracy, increasing $(\beta_1, \beta_2)_{n+1}$ in every iterations.

4.3. Algorithm

Variable splitting with continuation: Set $p = 0; (\beta_1)_0, (\beta_2)_0 > 0; \mathbf{R} = \mathbf{0}; \mathbf{S}_i = \mathbf{0}; i = 0, \ldots, q-1; \mathbf{\Gamma} = \mathbf{0}$

Repeat

Repeat

Update $\mathbf{\Gamma}$ by solving (14) using the CG scheme;

Shrinkage: $\mathbf{R} = \mathcal{S}_{\lambda_1/(\beta_1)_p}(\mathbf{\Gamma})$;

Shrinkage: $\mathbf{S}_i = \kappa_{\lambda_2/(\beta_2)_p}(\mathbf{\Gamma}); i = 0, \ldots, q-1$;

Until stopping criterion is satisfied.

$(\beta_1)_{p+1} = (\beta_1)_p * \mathrm{INC\_FACTOR}_1$;

$(\beta_2)_{p+1} = (\beta_2)_p * \mathrm{INC\_FACTOR}_2$;

$p = p + 1$

Until $\mathbf{R} \approx \mathbf{\Gamma}$ and $\mathbf{S}_i = \mathbf{\Phi}_i^H \mathbf{\Gamma} \mathbf{\Psi}_i; i = 0, \ldots, q-1$