Nuclear norm regularised dynamic mode decomposition

Shaobo Wang[1], Xiangyun Qing[2] ✉

[1]*Shanghai Environmental Protection Complete Engineering Co., Ltd., Shanghai Electric Group Co., Ltd., Shanghai, People's Republic of China*
[2]*Department of Automation, East China University of Science and Technology, Shanghai, People's Republic of China*
✉ *E-mail: xytsing@ecust.edu.cn*

Abstract: As a data-driven, equation-free decomposition method, the DMD can characterise dynamic behaviour of a non-linear system by using the DMD modes and eigenvalues. However, all current provable algorithms suffer from a separate procedure for obtaining the DMD modes and determining the number of modes. In this study, the authors propose a nuclear norm regularised DMD (NNR-DMD) algorithm that produces low-dimensional spatio-temporal modes. A nuclear norm regularisation term is added to the optimisation problem of the standard DMD algorithm for prompting the sparsity of the projected DMD modes. Split Bregman method is applied to solve the regularised convex, but non-smooth optimisation problem. Several numerical examples demonstrate the potential of the proposed NNR-DMD algorithm: (i) it can identify the low-dimensional spatio-temporal DMD modes in which each of them possesses a single temporal frequency; (ii) the reconstruction errors based on the sparse DMD modes can be reduced when it compares with the sparsity-promoting DMD algorithm penalising the l_1-norm of the vector of DMD amplitudes; and (iii) it can obtain low-dimensional coherent structures when the NNR-DMD algorithm is applied to coherency identification of generators in an interconnected power system.

1 Introduction

As a powerful tool developed in the fluid mechanics community, dynamic mode decomposition (DMD) can capture information about dynamics of a flow even when these dynamics are non-linear [1]. DMD modes obtained from data in simulations or in experiments identify coherent structures in fluid flows. Although the DMD was first introduced in [2] only for evaluating fluid dynamics, the DMD has been widely used to analyse observational data arising from complex, high-dimensional systems that may be described by a low-dimensional attractor subspace defined by a few coherent structures [3]. In [4], the DMD is used for background/foreground separation in videos. The DMD modes can be interpreted as stationary background pixels, or low-rank components of the data matrix. In [5], a DMD model is learned for behaviour-specific parameter configurations in physical human–robot interaction tasks. Therefore, the importance of the DMD is verified in the field of robotics. In [6], a physically-motivated DMD algorithm is presented to monitor the spatial and temporal dynamics of non-linear transient phenomena in power systems. The DMD modes can be used to identify the coherent generators.

One of the main advantages of using the DMD is that it is a data-driven, equation-free method to reconstruct the underlying low-dimensional dynamics of the system from snapshot measurements alone [7]. Although there are numerous data-driven, equation-free decomposition methods such as singular value decomposition (SVD), principal component analysis (PCA), independent component analysis, the DMD can characterise non-linear dynamics of the system by using the DMD modes and eigenvalues. Theoretically, the DMD can be interpreted as a numerical approximation to Koopman spectral analysis using linear transformations on Hilbert space to analyse Hamiltonian system [1]. For Koopman spectral analysis, a linear but infinite-dimensional Koopman operator is introduced to capture the full information of the nonlinear dynamical system even if the

governing dynamics of the system are finite dimensional. Thus, DMD is applicable to nonlinear systems since it approximates the Koopman modes and eigenvalues, but not the Koopman eigenfunctions [8]. For a comprehensive understanding of the connection between the DMD and other data-driven decomposition methods, we refer reader to [9].

The DMD method is closely related to proper orthogonal decomposition (POD) method which sometimes also called PCA or Karhunen–Loeve decomposition method introduced for the fluids community by Lumley [10] and Sirovich [11]. The POD method allows the given data to be represented as a linear combination of basic functions, or modes. The POD modes containing multi-frequency temporal components are spatial orthogonal. The DMD modes can be viewed as the linear combinations of the POD modes, which may be non-orthogonal. However, each DMD mode corresponds to a single-frequency component, which can provide an interpretation in terms of decay rate and frequency of this component. Therefore, although the DMD modes lack spatial orthogonality, they may be essential to capturing important system dynamics, while the POD modes fail intrinsically to provide any dynamical information. However, the major pitfall of the DMD method is that it is not entirely clear how to obtain to a mode set when a user only requires fewer modes. Consequently, a question of how best to extract fewer dominated modes can be urgent for some practical applications. For example, identifying dominant low frequency inter-area modes of oscillation has received increased importance for a near real-monitoring in a multi-area power system [6]. Another example is to identify a subset of DMD modes that retain the most important flow dynamic information for a fluid system. The first notable attempts toward this goal were made by Chen *et al.* in [12] with the introduction of optimised DMD (opt-DMD). For finding complex scalars to minimise the overall residual of fitting the given data set, a global optimisation technique that combined simulated annealing and the Nelder–Mead simplex method was employed, which conducts an intractable combinatorial search. In

[13, 14], a new method, called optimal mode decomposition (OMD), was presented to estimate the linear dynamics of a high-dimensional system. Both the dynamic matrix and the low-rank basis were simultaneously optimised by employing a gradient-based algorithm. However, the algorithm is computationally expensive and the rank should be firstly fixed. Recently, a sparsity-promoting DMD (SP-DMD) algorithm has been developed to balance between the quality of approximation and the number of modes in [15]. The method consists of two steps, in which the first step seeks a sparsity structure and the second step determines the optimal the non-zero amplitudes associating with the sparsity structure. The method can be viewed as a convex relation of optimisation problem in the opt-DMD algorithm by using l_1 norm. However, the SP-DMD method cannot optimise over both the modes and the number of modes simultaneously because the modes are fixed.

In this paper, we develop a nuclear norm regularised DMD (NNR-DMD) algorithm. This algorithm picks the appropriate number of modes or model complexity by adding an additional nuclear norm term into the optimisation problem of the DMD algorithm in the least-square sense. The objective of the proposed algorithm is straightforwardly to determine a low-rank representation of the matrix that captures the dynamic inherent in the data sequence. Unfortunately, rank minimisation is NP-hard in general. Thus, nuclear norm of the matrix instead of its rank, as a popular convex heuristic [16], is used to obtain a regularised version of the optimisation problem. Recently, the nuclear norm regularisation approach has been successfully applied in matrix factorisation based low-rank representation in [17–19]. Finally, split Bregman method is then employed to solve the resulting regularised optimisation problem with a non-smooth nuclear norm term and a smooth Frobenius norm term. The split Bregman method, proposed in [20, 21], is admirably suitable for dealing with the decoupling between the nuclear and Frobenius norm terms in the iterations. In each iteration of the proposed NNR-DMD algorithm, the optimisation problem can be decomposed into some easily solvable subproblems which are computationally fast and easy to code.

The rest of the paper is organised as follows. We first present an algorithmic definition of the SVD-based DMD, its different interpretations and sparsity prompting DMD algorithms in Section 2. The proposed NNR-DMD algorithm is then presented in Section 3. Section 4 provides several numerical test examples. Concluding remarks are drawn in Section 5.

Notations used in this paper are shown as follows: $\mathbb{C}^n$ denotes the n-dimensional complex space. The set of all $n \times m$ matrices with complex entries is denoted by $\mathbb{C}^{n \times m}$. For any $A \in \mathbb{C}^{n \times m}$, A^* denotes the complex-conjugate-transpose of A, $\|A\|_{\text{fro}}$ denotes the Frobenius norm of A, $\|A\|_*$ denotes the nuclear norm of A, which is the sum of all singular values of A. diag(x) is a diagonal matrix with its diagonal entries being entries of a vector x. $\Re e\{\cdot\}$ and $\Im m\{\cdot\}$ denote the real and the imaginary parts of a complex quantity, respectively. Finally, we denote the identity matrix by I whose dimension should be clear from the context.

2 Dynamic model decomposition

2.1 Preliminaries and problem statement

We assume that a sequential set of data vectors $\{x_0, x_1, \ldots, x_m\}$ is generated by a linear system

$$x_{k+1} = A x_k \qquad (1)$$

where $x_k \in \mathbb{C}^n$ is the kth snapshot, $A \in \mathbb{C}^{n \times n}$ is an unknown (time-independent) matrix. Obviously, the evolution of x is governed by the eigenvalues of A for the linear system. When the data set are generated from a non-linear system, it is assumed that A is an operator to approximate the dynamics of the non-linear system.

The computation of DMD modes proceeds as follows:

Algorithm 1 (standard DMD):

1. The data are grouped into matrices

$$\begin{aligned} X_0 &= [x_0, x_1, \ldots, x_{m-1}] \\ X_1 &= [x_1, x_2, \ldots, x_m] \end{aligned} \qquad (2)$$

2. Compute the reduced SVD of X_0

$$X_0 = U \Sigma V^* \qquad (3)$$

where U is $n \times r$, Σ is diagonal and $r \times r$, V is $m \times r$, and r is the rank of X_0.

3. Assume that an optimal representation F of the matrix A in the basis spanned by the POD modes of X_0 is provided by the DMD algorithm

$$A \simeq U F U^* \qquad (4)$$

Compute the least-squares fit A that satisfies $X_1 = A X_0$ by minimising the Frobenius norm of the difference between X_1 and X_0

$$\min_{F} \|X_1 - U F \Sigma V^*\|_{\text{fro}}^2 \qquad (5)$$

where F is $r \times r$. The optimal solution to (5) is obtained by

$$F_{\text{dmd}} = U^* X_1 V \Sigma^{-1} \qquad (6)$$

4. Compute the eigenvalues and eigenvectors of F_{dmd}

$$F_{\text{dmd}} = W \Lambda W^{-1} \qquad (7)$$

where $\Lambda = \text{diag}([\lambda_1 \ \lambda_2 \ \ldots \ \lambda_r]) \in \mathbb{C}^{r \times r}$ is a diagonal matrix consisting of eigenvalues λ_j and $W = [\omega_1 \ \omega_2 \ \ldots \ \omega_r] \in \mathbb{C}^{r \times r}$ is the matrix of right eigenvectors, respectively.

5. Compute the DMD modes

$$\Phi = U W \qquad (8)$$

where $\Phi = [\phi_1 \ \phi_2 \ \ldots \ \phi_n]^{\text{T}} \in \mathbb{C}^{n \times r}$ is the matrix of projected DMD modes.

The motivation of the DMD algorithm can be interpreted using linear dynamical system theory in the noise-free case [15]. Suppose the dynamics of a linear system of extremely high observational dimension are governed by a low-dimensional state space as

$$\theta_{t+1} = F_{\text{dmd}} \theta_t \qquad (9)$$

The eigenvalues and eigenvectors of F_{dmd} determine the dynamic behaviour of the linear system. Thus, each state vector θ_t can be determined by

$$\theta_t = F_{\text{dmd}}^t \theta_0 = W \Lambda^t W^* \theta_0 = \sum_{i=1}^{r} \omega_i \lambda_i^t \omega_i^* \theta_0 = \sum_{i=1}^{r} \omega_i \lambda_i^t \alpha_i \qquad (10)$$

where each $\alpha_i = \omega_i^* \theta_0$ represents the ith modal of the initial condition θ_0.

The state vector θ_t is mapped into a higher dimensional observational vector by

$$x_t \simeq U \theta_t \qquad (11)$$

Thus, we have

$$x_{t+1} \simeq U\theta_{t+1} = UF_{\mathrm{dmd}}\theta_t = \underbrace{UF_{\mathrm{dmd}}U^*}_{A} \underbrace{U\theta_t}_{x_t} \simeq Ax_t \qquad (12)$$

which is an approximate of a linear dynamic system as (1).

Furthermore, experimental or numerical snapshots can be approximated by using a linear combination of the DMD modes as follows

$$x_t \simeq \sum_{i=1}^{r} \phi_i \lambda_i^t \alpha_i, \quad t \in \{0, 1, \ldots, m-1\} \qquad (13)$$

The unknown vector $\alpha = [\alpha_1, \ldots, \alpha_r]^{\mathrm{T}}$ can be determined by solving the following optimisation problem

$$\min_{\alpha} \left\| X_0 - \Phi D_\alpha V_{\mathrm{and}} \right\|_{\mathrm{fro}}^2 \qquad (14)$$

where $D_\alpha \triangleq \mathrm{diag}(\alpha)$ and V_{and} is a Vandermode matrix written as

$$V_{\mathrm{and}} \triangleq \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_r & \cdots & \lambda_r^{m-1} \end{bmatrix} \qquad (15)$$

In [6], the DMD has been interpreted as characterising global dynamic behaviour of spatio-temporal data. The jth element of the vector x_t denotes an observation at the j measurement point at the time t. By substituting (7) into (6), we have

$$U^*X_1 V\Sigma^{-1} = W\Lambda W^{-1} \qquad (16)$$

Thus, the data sequence X_1 can be expressed as

$$X_1 \simeq \underbrace{UW}_{\text{Spatial structure}} \Lambda \underbrace{W^{-1}\Sigma V^*}_{\text{Temporal structure}} = \Phi\Lambda\Gamma(t) \qquad (17)$$

where each row of $\Gamma(t) \triangleq W^{-1}\Sigma V^* \in \mathbb{C}^{r \times (m-1)}$ represents a temporal pattern containing a single-frequency component, and each row of Φ represents a spatial mode (DMD mode). The dynamical modes, $\phi_j = U\omega_j$, represent the spatial coherent structure. Moreover, the decay rating $|\lambda_j|$ associating with the jth mode is related to the damping in case of sampled dynamics, while $f_j \triangleq \Im m\{\log(\lambda_j)\}/(2\pi\Delta t)$ is the frequency of the jth temporal pattern.

2.2 Sparsity-promoting DMD

For the OMD algorithm proposed in [13, 14], a low-rank approximation of the flow dynamics is obtained by solving a rank-constrained matrix optimisation problem as follows

$$\begin{aligned} \min_{A} &\ \left\| X_1 - AX_0 \right\|_{\mathrm{fro}}^2 \\ \text{s.t.} &\ A \triangleq \Phi F \Phi^{\mathrm{T}} \\ &\ F \in \mathbb{C}^{r \times r}, \Phi \in \mathcal{S}_{n,r} \end{aligned} \qquad (18)$$

where $\mathcal{S}_{n,r}$ is the set of Stiefel matrices

$$\mathcal{S}_{n,r} = \{\Phi \in \mathbb{C}^{n \times r} | \Phi^{\mathrm{T}}\Phi = I, r \ll n\} \qquad (19)$$

Although this optimisation problem can be solved using a technique based on optimisation on the Grassman manifold, it inherits the drawbacks of non-convex problems such as local convergence and being sensitive to initial solutions.

For the SP-DMD algorithm proposed in [15]. A sparsity structure is sought by solving a l_1-norm regularised optimisation problem in its first step as follows

$$\min_{\alpha} \left\| X_0 - \Phi D_\alpha V_{\mathrm{and}} \right\|_{\mathrm{fro}}^2 + \gamma \sum_{i=1}^{r} |\alpha_i| \qquad (20)$$

The sparsity structure is identified because the l_1-norm can prompt the sparsity of the vector α.

In its second step, the optimal values of the non-zeros amplitudes are determined by solving a constrained optimisation problem only associating with the support of the vector α.

Although the SP-DMD algorithm can achieve a global optimal solution by solving two convex optimisation problems, it assumes that the modes are fixed in advance. Its task is to find the optimal and sparse modes from the given mode set. However, the modes and their corresponding amplitudes are interacted on each other. Thus, we seek a low-rank matrix straightforwardly to obtain sparsity-promoting DMD.

3 Nuclear norm regularised DMD

3.1 Problem formulations

Notice that the optimisation (5) can be equivalently written as

$$\min_{F} \left\| U^*X_1 - F\Sigma V^* \right\|_{\mathrm{fro}}^2 \qquad (21)$$

By defining $Y = U^*X_1$ and $X = \Sigma V^*$, the DMD can be generalised to a matrix optimisation problem as follows

$$\min_{F} \| Y - FX \|_{\mathrm{fro}}^2 \qquad (22)$$

We approach the sparseness-inducing regularisation for DMD by augmenting the objective function in (22) with an addition term, rank(F), that penalises the rank of the matrix F

$$\min_{F} \frac{1}{2} \| Y - FX \|_{\mathrm{fro}}^2 + \mu \cdot \mathrm{rank}(F) \qquad (23)$$

where μ is a positive regularisation parameter balancing the least-square fitting term and the low-dimensional representation of the matrix F. However, as directly minimising the rank is NP-hard in general. Thus, paralleling the use of the l_1-norm in sparse approximation of cardinality minimisation of a vector, a convex relation for rank minimisation is adopted by replacing the rank minimisation by nuclear norm minimisation. The resulting optimisation problem can be written as follows

$$\min_{F} \frac{1}{2} \| Y - FX \|_{\mathrm{fro}}^2 + \mu \| F \|_* \qquad (24)$$

Then, the modes Φ can be achieved according the optimal solution F_{dmd} of (24). Computing the eigenvalue decomposition $F_{\mathrm{dmd}} = W\Lambda W^{-1}$, we have $\Phi = U\hat{W}$ where $\hat{W}$ is a matrix consisting of the eigenvectors with non-zeros eigenvalues. Owing to the introduced nuclear norm to prompt the low-rank representation of F_{dmd}, the number of non-zeros eigenvalues of F_{dmd} may less far than the rank r.

Defining $\mathcal{F}: \mathbb{C}^{r \times m} \to \mathbb{C}^{r^2}$ as a linear operator as

$$\mathcal{X}(F) = \hat{X}\mathrm{vec}(F) \qquad (25)$$

where vec(F) creates a column vector from the matrix F by stacking the column of F and $\hat{X}$ is the Kronecker product of a $r \times r$ identity matrix and an $m \times r$ matrix X^{T}, the optimisation problem (24) can

be equivalently written as

$$\min_{F} \frac{1}{2} \left\| \mathrm{vec}(Y) - \mathcal{X}(\boldsymbol{F}) \right\|_2^2 + \mu \|\boldsymbol{F}\|_* \qquad (26)$$

This convex program is an over-determined problem as $r^2 \leq r \times m$. Unlike recovery of low-rank matrix from nuclear norm minimisation in the field of compressed sensing, the regularisation using nuclear norm minimisation is to avoid over-fitting and seek a 'simplest' model.

3.2 Split Bregman iteration for nuclear norm regularised DMD

We reformulate (24) by introducing a new variable H and a constraint $\boldsymbol{F} = H$ as

$$\begin{aligned} \min_{\boldsymbol{F},H} \ & \frac{1}{2} \|Y - \boldsymbol{F}X\|_{\mathrm{fro}^2} + \mu \|H\|_* \\ \text{s.t.} \quad & \boldsymbol{F} = H \end{aligned} \qquad (27)$$

The solving convex programming problem with separable structure is easy to produce an algorithm that consists of simple, ease to compute steps. Then, (27) can be converted into an unconstrained problem by use of penalty technique

$$\min_{\boldsymbol{F},H} \frac{1}{2} \|Y - \boldsymbol{F}X\|_{\mathrm{fro}}^2 + \mu \|H\|_* + \frac{\eta}{2} \|H - \boldsymbol{F}\|_{\mathrm{fro}}^2 \qquad (28)$$

Thus, we apply the split Bregman method and get the split Bregman iteration as follows

$$\begin{cases} (\boldsymbol{F}^{k+1}, H^{k+1}) = \underset{\boldsymbol{F},H}{\mathrm{argmin}} \ \frac{1}{2} \|Y - \boldsymbol{F}X\|_{\mathrm{fro}}^2 \\ \qquad + \mu \|H\|_* + \frac{\eta}{2} \|H - \boldsymbol{F} - B^k\|_{\mathrm{fro}}^2 \\ B^{k+1} = B^k + (\boldsymbol{F}^{k+1} - H^{k+1}) \end{cases} \qquad (29)$$

The first minimisation of (29) can be split into two simpler subproblems as

$$\boldsymbol{F}^{k+1} = \underset{\boldsymbol{F}}{\mathrm{argmin}} \ \frac{1}{2} \|Y - \boldsymbol{F}X\|_{\mathrm{fro}}^2 + \frac{\eta}{2} \|H^k - \boldsymbol{F} - B^k\|_{\mathrm{fro}}^2 \qquad (30)$$

$$H^{k+1} = \underset{H}{\mathrm{argmin}} \ \mu \|H\|_* + \frac{\eta}{2} \|H - \boldsymbol{F}^{k+1} - B^k\|_{\mathrm{fro}}^2 \qquad (31)$$

Equation (30) is convex and differentiable, which can obtain a close form solution as follows

$$\boldsymbol{F}^{k+1} = [YX^* + \eta(H^k - B^k)](XX^* + \eta I)^{-1} \qquad (32)$$

Equation (31) can be solved by singular value soft-thresholding method performing the shrinkage operation on singular values of H [22]. Let the matrix $\boldsymbol{Z} = \boldsymbol{F}^{k+1} + B^k$ has SVD $\boldsymbol{Z} = U_H \hat{\Sigma} V_H^{\mathrm{T}}$. Then the solution of the optimisation problem (31) is given by

$$H^{k+1} = U_H S_{\mu/\eta}[\hat{\boldsymbol{\Sigma}}] V_H^{\mathrm{T}} \qquad (33)$$

where $S_{\mu/\eta}[\hat{\boldsymbol{\Sigma}}]$ is the shrinkage operation applied entrywise to the matrix $\hat{\boldsymbol{\Sigma}}$ as

$$(S_{\mu/\eta}[\hat{\boldsymbol{\Sigma}}])_{ij} = \max\left(0, \left|\hat{\boldsymbol{\Sigma}}_{ij}\right| - \mu/\eta\right) \hat{\boldsymbol{\Sigma}}_{ij} / \left|\hat{\boldsymbol{\Sigma}}_{ij}\right| \qquad (34)$$

and $|\cdot|$ is the modulus of the complex matrix entry.

Now, we present complete algorithm for NNR-DMD as Algorithm 2 (see Fig. 1).

Two stop criteria are adopted. The first criterion is

$$\frac{\|F^k - F^{k-1}\|_{\mathrm{fro}}}{\|F^{k-1}\|_{\mathrm{fro}}} \leq 10^{-5}.$$

The other criterion is that the maximum number of iterations is set to 500.

The convergence property of the proposed algorithm NNR-DMD is shown in the following theorem, which is similar to that in the work [23, 24].

Theorem 1: Assume that there exists at least one solution $\hat{F}$ of (24). Assume $\mu > 0$. Then, the following property for the split Bregman iteration in Algorithm 2 (NNR-DMD) holds

$$\lim_{k \to +\infty} \frac{1}{2} \|Y - F^k X\|_{\mathrm{fro}}^2 + \mu \|F^k\|_* = \frac{1}{2} \|Y - \hat{F} X\|_{\mathrm{fro}}^2 + \mu \|\hat{F}\|_* \qquad (35)$$

3.3 Complexity analysis

In this part, we discuss the computational complexity of the proposed algorithm NNR-DMD. In (3), the computational complexity of performing SVD of X_0 is $O(nm^2)$; The computational complexity of matrix multiplication for F^{k+1} is $O(r^3)$; The computational complexity of SVD on the small matrix Z is $O(r^3)$. The computational complexity of matrix multiplication for H^{k+1} is also $O(r^3)$. Therefore, the total computational complexity of the proposed algorithm is $O(nm^2 + 3kr^3)$, where k is the number of iterations. As the rank r of the matrix X_0 satisfies $r \leq \min(m, n)$ and the number m of the snapshots can be picked to a relatively small one, the computational complexity of the proposed algorithm can be acceptable for many signal processing tasks.

4 Numerical examples

4.1 Identifying the temporal frequency

In this example, the proposed algorithm (NNR-DMD) is to identify the temporal frequency from the spatio-temporal data generating from a synthetic sinusoidal flow of the form

$$f(x, t) = \sin(kx - 2\pi f t)\, e^{\gamma t} \qquad (36)$$

with parameters $\gamma = 1$, $k = 4$, and $f = 2$. The flow was used in [13, 25]. The number of temporal snapshots was $m = 50$, taken at time intervals $dt = \pi/50$. The number of spatial samples was $n = 201$ by taking at sampling intervals $dx = \pi/50$. The resulting data $\{x_0, x_1, \ldots, x_m\}$ was corrupted by adding zeros-mean Gaussian

Algorithm 2

Initialization: $B^0 = I; H^0 = 0$
while "none of stop criteria is met" do
    $F^{k+1} = [YX^* + \eta(H^k - B^k)](XX^* + \eta I)^{-1}$
    $Z = F^{k+1} + B^k$ where $Z = U_H \hat{\Sigma} V_H^T$
    $H^{k+1} = U_H S_{\mu/\eta}[\hat{\Sigma}] V_H^T$
    $B^{k+1} = B^k + (F^{k+1} - H^{k+1})$
    $k \leftarrow k + 1$
end while

Fig. 1 *Algorithm 2 (NNR-DMD)*

noise with covariance $\sigma = 0.1$. The parameters of the NNR-DMD were set as $\mu = 10$ and $\eta = 100$.

The eigenvalues and frequencies on 250 data ensembles using Monte Carlo sampling method at the covariance level were averaged. We obtained only a pair of conjugate complex eigenvalues by using the NNR-DMD algorithm. The corresponding positive frequency was 1.9997. This means that the single frequency is correctly identified by applying the NNR-DMD algorithm. The temporal pattern only contains a single-frequency component, just as the assumption holds.

We provide the comparison of the running time of the proposed NNR-DMD algorithm and the OMD algorithm in [13]. The experiment was performed with Matlab 2014a on an Intel(R) Core (TM) i7-4700MQ CPU 2.39 GHz Laptop running Windows8 with 4 GB main memory. The running time of the proposed NNR-DMD algorithm is less than 0.8 s, while the OMD algorithm takes around 3.8 s.

4.2 Reducing the reconstruction errors

The data set of two-dimensional Poiseuille flow provided by Jovanovic *et al.* [15] was used to analyse the reconstruction errors. The performance of the NNR-DMD algorithm is compared with the SP-DMD algorithm given in [15]. The data set was collected from 150 Chebyshev collocation points and 100 temporal snapshots. Thus, we have $n = 150$ and $m = 50$. The rank of the matrix of snapshots X_0 is $r = 26$. The parameters of the NNR-DMD algorithms for this example were set as $\mu = 1$ and $\eta = 10$.

For the NNR-DMD algorithm, the unknown vector of amplitudes α was obtained by solving a convex optimisation problem

$$\min_{\alpha} \ J(\alpha) = \left\| X_0 - \Phi D_\alpha V_{\text{and}} \right\|_{\text{fro}}^2 \tag{37}$$

Here, Φ and V_{and} were obtained by applying the NNR-DMD algorithm or the standard DMD algorithm.

A relative residual value was introduced to compare the reconstruction performance, which is defined as

$$\%\Pi_{\text{loss}} \triangleq 100 \sqrt{\frac{J(\alpha)}{J(0)}} = 100 \frac{\left\| X_0 - \Phi D_\alpha V_{\text{and}} \right\|_{\text{fro}}}{\left\| X_0 \right\|_{\text{fro}}} \tag{38}$$

For the NNR-DMD algorithm, five DMD modes were obtained and the relative residual value for the optimal vector of amplitudes was 2.75%. According the results of [15], the reasonable number of modes in the Poiseuille flow may be achieved with six modes. It means that the number of modes using our algorithm approximates the ideal result. However, the relative residual value resulting from the SP-DMD algorithm exceeded to 8%, which shows significant performance deterioration compared with our algorithm. The main reason lies in that our algorithm optimises the number of modes and DMD modes Φ simultaneously, whereas the SP-DMD algorithm only find the best modes from the mode set obtained by the standard DMD algorithm.

We also provide the comparison of the running time of the proposed NNR-DMD algorithm and the SP-DMD algorithm. The experiment was performed on the same test environment as the previous experiment. The running time of the proposed NNR-DMD algorithm is less than 0.4 s, whereas the SP-DMD algorithm takes around 1.2 s.

4.3 Identifying coherent generators in a power system

In interconnected power systems, coherency identification of generators can be very important for reducing the dynamic model order and initiating islanding when large changes in loads or generation may lead the power system into different balanced islands to avoid wide-area blackouts [26, 27]. Intuitively, a group of generators is said to be coherent if after some disturbances, they present similar time-domain response trends. Recently, data-driven

methods using online measurements and signal processing techniques such as PCA, ICA, and Hilbert–Huang transform [28–30], have received considerable attention because the methods are free of very accurate model of the study system. In particular, methods borrowing from the fluid flow field have also been applied to identify the coherency generators. In [31], a systematic and rigorous way for coherency identification based on the Koopman modes has been presented. However, the Koopman modal expansions for the high dimensionality are computationally challengeable. Thus, in [6], the DMD algorithm has been proposed for large datasets.

We performed coherency identification in the New England 39-bus test system, which is shown in Fig. 2. Non-linear simulations of the test system were performed using the MATLAB-based PST toolbox [32]. The system was being simulated for 5 s and lines 16 and 17 were tripped at $t = 1$ s. Measurements were recorded over 5 s at a rate 50 samples per second. The speed deviations of all ten generators (G1–G10) in the 5 s time interval are shown in Fig. 3. We selected the measurements from the first second to fifth second to construct the snapshot matrices X_0 and X_1. The generator 1 was set as the reference generator. The speed deviations after subtracting the mean values were used for the following experiments. The fast Fourier transform (FFT) of the speed deviations is shown in Fig. 4. From Fig. 4, it can be seen that their dominant frequency is about 0.59 Hz. By using the small-signal analysis method provided in [32, 33], the dominant low frequency inter-area modes of oscillation are about 0.58 Hz, which coincide with the results of time transient analysis.

The μ in the NNR-DMD algorithm was set to 0.001 because the speed deviations after subtracting the mean values are small. The parameter η was set to 0.01. The eigenvalues λ_is and their corresponding frequencies by applying the DMD algorithm and the proposed algorithm NNR-DMD are shown in Table 1. Fig. 5 also shows a plot of the eigenvalues λ_is. As seen in Table 1 and Fig. 4, the NNR-DMD can obtain only a dominant frequency 0.5805 Hz, whereas the DMD algorithm exhibits all frequencies. However, the dominant frequency 0.5805 Hz obtained using the NNR-DMD algorithm is closer to the theoretical result than the frequency 0.5477 Hz obtained using the DMD algorithm.

Clusters of coherent generators were identified from the spatial modes corresponding to the dominant low frequencies 0.54 and 0.58 Hz for the DMD and NNR-DMD algorithms are shown in Figs. 6 and 7, respectively. The eighth column of matrix Φ for the DMD algorithm shows the strongest clustering information,
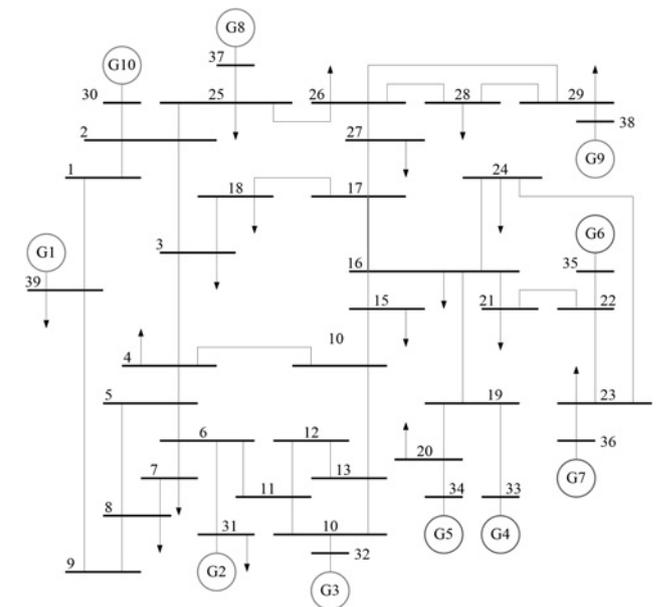
Fig. 2 *New England 39-bus test system model*
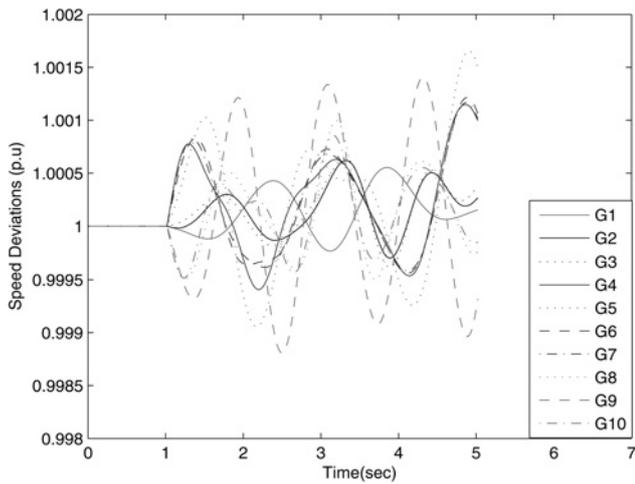
Fig. 3 Generator speed deviations

Fig. 4 Spectra analysis results for speed deviations of generators from G1 to G10

whereas the second column of matrix **Φ** for the NNR-DMD algorithm does. By setting the number of groups to two and using the Coherency Toolbox based on the slow coherency theory [32, 33], a benchmark of coherency identification of two coherent generator groups, $\{1, 2, 3, 8, 9, 10\}$ and $\{4, 5, 6, 7\}$, was obtained. As seen in Figs. 6 and 7, although the generator clustering results of the two algorithms are identical by using the signatures of real modes and match the benchmark results, the NNR-DMD

Table 1 Comparisons of eigenvalues λ_is and their associated frequencies

DMD		NNRDMD	
λ_i	Frequency	λ_i	Frequency
1.0480	0	1.4078	0
$0.9877 + 0.1660i$	1.3247	$0.9350 + 0.0683i$	0.5805
$0.9877 - 0.1660i$	−1.3247	$0.9350 - 0.0683i$	−0.5805
$0.9951 + 0.1038i$	0.8267	0.4902	0
$0.9951 - 0.1038i$	−0.8267	0.1662	0
$0.9900 + 0.1382i$	1.1039	0.4475	0
$0.9900 - 0.1382i$	−1.1039	–	–
$0.9948 + 0.0686i$	0.5477	–	–
$0.9948 - 0.0686i$	−0.5477	–	–
0.0084	0	–	–

Fig. 5 Eigenvalues λ_is obtained from the DMD and NNR-DMD algorithms

algorithm can find the dominant mode with the strongest impact on the spatial coherent structure by using the nuclear norm regularisation, whereas the DMD algorithm provides more modes needed to be further selected for coherency identification.

5 Concluding remarks

This paper has introduced a NNR-DMD algorithm, as an extension of the standard DMD algorithm, for identifying a low-rank DMD matrix. The low-rank matrix is obtained by introducing a nuclear norm term of the matrix to the stand least-squares optimisation problem. The resulting regularised version of the optimisation problem is a convex, but non-smooth optimisation problem which the split Bregman algorithm provides an effective tool for decoupling the variables with different norms and updates the variables alternately. We have shown that the solving process is computationally efficient and easy to implement. The numerical results show that the proposed NNR-DMD algorithm can be efficient to identify the temporal frequency from the spatio-temporal data, reduce the reconstruction errors by comparing with the OMD and SP-DMD algorithms. The algorithm is also suitable for coherency identification of generators in an interconnected power system, which is verified in the England 39-bus test system.
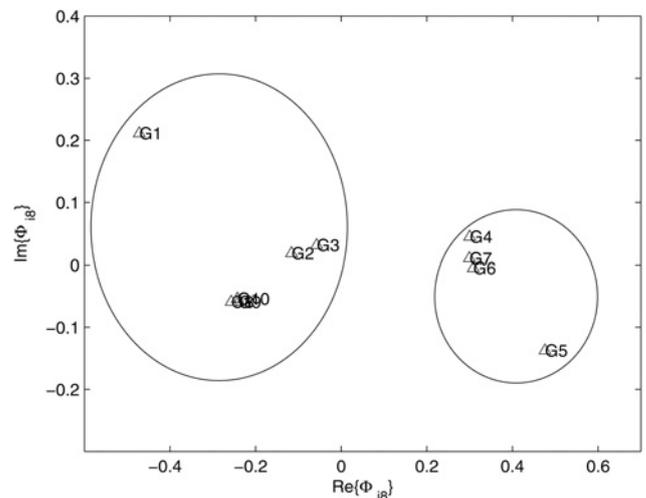
Fig. 6 Coherency identification using the DMD algorithm
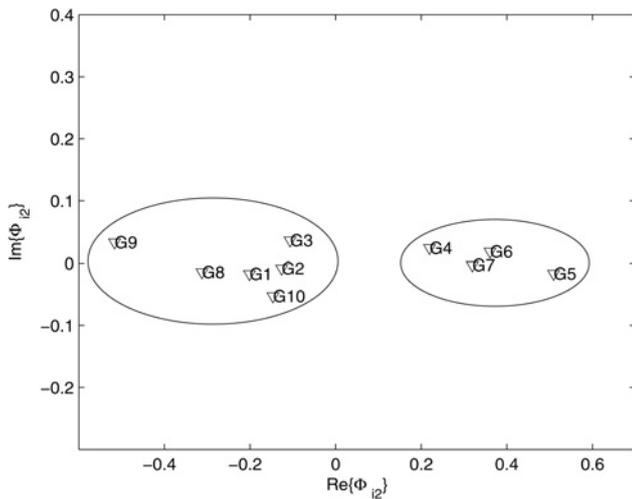
Fig. 7 *Coherency identification using the NNR-DMD algorithm*

6 References

1 Rowley, C.W., Mezic, I., Bagheri, S., *et al.*: 'Spectral analysis of nonlinear flows', *J. Fluid Mech.*, 2009, **641**, pp. 115–127

2 Schmid, P.J., Sesterhenn, J.: 'Dynamic mode decomposition of numerical and experimental data'. 61th Annual Meeting of the APS Division of Fluid Dynamics, American Physical Society, November 2008

3 Brunton, S.L., Proctor, J.L., Kutz, J.N.: 'Compressive sampling and dynamic mode decomposition'. arXiv preprint http://arxiv.org/abs/1312.5186, accessed January 2016

4 Grosek, J., Kutz, J.N.: 'Dynamic mode decomposition for real-time background/foreground separation in video'. arXiv preprint http://arxiv.org/abs/1404.7592, accessed January 2016

5 Berger, E., Sastuba, M., Vogt, D., *et al.*: 'Estimation of perturbations in robotic behavior using dynamic mode decomposition', *Adv. Robot.*, 2015, **29**, (5), pp. 331–343

6 Barocio, E., Pal, B.C., Thornhill, N.F., *et al.*: 'A dynamic mode decomposition framework for global power system oscillation analysis', *IEEE Trans. on Power Syst*, 2015, **30**, (6), pp. 2902–2912

7 Proctor, J.L., Brunton, S.L., Kutz, J.N.: 'Dynamic mode decomposition with control'. arXiv preprint http://arxiv.org/abs/1409.6358, accessed January 2016

8 Williams, M.O., Kevrekidis, I.G., Rowley, C.W.: 'A data-driven approximation of the Koopman operator: extending dynamic mode decomposition', *J. Nonlinear Sci.*, 2015, **25**, (6), pp. 1307–1346

9 Tu, J.H., Rowley, C.W., Luchtenburg, D.M.: 'On dynamic mode decomposition: theory and applications', *J. Comput. Dyn.*, 2014, **1**, (2), pp. 391–421

10 Lumley, J.L.: 'Stochastic tools in turbulence' (Academic, 1970)

11 Sirovich, L.: 'Turbulence and the dynamics of coherent structures', *Q. Appl. Maths.*, 1987, **45**, pp. 561–590

12 Chen, K.K., Tu, J.H., Rowley, C.W.: 'Variants of dynamic mode decomposition: boundary condition, Loopman, and Fourier analyses', *J. Nonlinear Sci.*, 2012, **22**, (6), pp. 887–915

13 Goulart, P.J., Wynn, A., Pearson, D.: 'Optimal mode decomposition for high dimensional systems'. Proc. 51st IEEE Conf. on Decision and Control, 2012, pp. 4965–4970

14 Wynn, A., Pearson, D.S., Ganapathisubramani, B., *et al.*: 'Optimal mode decomposition for unsteady flows', *J. Fluid Mech.*, 2013, **733**, pp. 473–503

15 Jovanovic, M.R., Schmid, P.J., Nichols, J.W.: 'Sparsity-promoting dynamic mode decomposition', *Phys. Fluids*, 2014, **26**, pp. 024103-1–024103-22

16 Fazel, M.: 'Matrix rank minimization with applications'. PhD thesis, Standford University, 2002

17 Liu, Y., Jiao, L.C., Shang, F., *et al.*: 'An efficient matrix bi-factorization alterative optimization method for low-rank matrix recovery and completion', *Neural Netw.*, 2013, **48**, pp. 8–18

18 Liu, Y., Jiao, L.C., Shang, F.: 'An efficient matrix factorization based low-rank representation for subspacing clustering', *Pattern Recognit.*, 2013, **46**, pp. 284–292

19 Shang, F., Liu, Y., Tong, H., *et al.*: 'Robust bilinear factorization with missing and grossly corrupted observations', *Information Sciences*, 2015, **307**, pp. 53–72

20 Goldstein, T., Osher, S.: 'The split Bregman method for l1-regularized problem', *SIAM J. Imaging Sci.*, 2009, **2**, (2), pp. 323–343

21 Cai, J.F., Osher, S., Shen, Z.: 'Split Bregman methods and frame based image restoration', *Mutiscale Model. Simul.*, 2009, **8**, (2), pp. 337–369

22 Cai, J.F., Candes, E.J., Shen, Z.W.: 'A singular value thresholding algorithm for matrix completion', *SIAM Journal of Optimization*, 2010, **20**, (4), pp. 1956–1982

23 Yin, W., Osher, S., Goldfarb, D., *et al.*: 'Bregman iterative algorithms for l_1-minimization with applications to compressed sensing', *SIAM J. Imaging Sci.*, 2008, **1**, (1), pp. 143–168

24 Zou, J., Fu, Y., Zhang, Q., *et al.*: 'Split Bregman algorithm for multiple measurement vector problem', *Multidim Syst. Sign. Process.*, 2015, **26**, pp. 207–224

25 Duke, D., Soria, J., Honnery, D.: 'An error analysis of the dynamic mode decomposition', *Exp. Fluids*, 2011, **55**, (2), pp. 529–542

26 Xu, G., Vittal, V.: 'Slow coherency based cutset determination algorithm for large power system', *IEEE Trans. Power Syst.*, 2010, **25**, (2), pp. 877–884

27 Qing, X.: 'Controlled islanding of power systems using label propagation', *IEEJ Trans. Electric. Electron. Eng.*, 2015, **10**, (3), pp. 256–261

28 Anaparthi, K.K., Chaudhuri, B., Thronhill, N.F., *et al.*: 'Coherency identification in power systems through principal component analysis', *IEEE Trans. Power Syst.*, 2005, **20**, (3), pp. 1658–1660

29 Ariff, M.A.M., Pal, B.C.: 'Coherency identification in interconnected power system – an independent component analysis approach', *IEEE Trans. Power Syst.*, 2013, **28**, (2), pp. 1747–1755

30 Senroy, N.: 'Generator coherency using the Hilbert-Huang transform', *IEEE Trans. Power Syst.*, 2008, **23**, (4), pp. 1701–1708

31 Susuki, Y., Mezic, I.: 'Nonlinear Koopman modes and power stability assessment without models', *IEEE Trans. Power Syst.*, 2014, **29**, (2), pp. 899–907

32 Chow, J.H., Cheung, K.W.: 'A toolbox for power system dynamics and control engineering education and research', *IEEE Trans. Power Syst.*, 1992, **7**, (4), pp. 1559–1564

33 Chow, J.H.: 'Power system coherency and model reduction' (Springer, 2013)