

# Total-Variation Mode Decomposition<sup>\*</sup>

Ido Cohen, Tom Berkov, and Guy Gilboa

Technion - Israel Institute of Technology, Haifa 3200003, Israel  
{idoc, ptom}@campus.technion.ac.il, guy.gilboa@ee.technion.ac.il

**Abstract.** In this work we analyze the *Total Variation* (TV) flow applied to one dimensional signals. We formulate a relation between *Dynamic Mode Decomposition* (DMD), a dimensionality reduction method based on the Koopman operator, and the spectral TV decomposition. DMD is adapted by time rescaling to fit linearly decaying processes, such as the TV flow. For the flow with finite subgradient transitions, a closed form solution of the rescaled DMD is formulated. In addition, a solution to the TV-flow is presented, which relies only on the initial condition and its corresponding subgradient. A very fast numerical algorithm is obtained which solves the entire flow by elementary subgradient updates.

**Keywords:** *Total Variation*-flow · *Total Variation*-spectral decomposition · *Dynamic Mode Decomposition* · Time reparametrization.

## 1 Introduction

Finding latent information in high dimensional data is one of the most challenging tasks in data analysis. Often, dimensionality reduction techniques are used to reveal the latent modes. In this work, we bridge between two methods, *Total Variation* (TV) spectral decomposition [13] and *Dynamic Mode Decomposition* (DMD) [17]. These methods analyze gradient descent flows, represented as PDEs, and allow to perform filtering, spectrum analysis, and signal reconstruction.

TV-flow [1], the steepest descent process with respect to the TV functional, decays piecewise linearly and has finite support in time. This was used to compute the spectral TV components and to analyze their properties, see [12,13,6,14,5]. The two attributes are at odds with *Dynamic Mode Decomposition* (DMD), an approximation of the Koopman operator [16], which is described as an exponential data fitting algorithm [2]. Moreover, it was shown in [10] that DMD converges to highly inaccurate solutions, in certain cases, for flows derived by a  $\gamma$ -homogeneous operator when  $\gamma \neq 1$ . One way to improve the accuracy of DMD is by adding nonlinear measurements to the state space [20],[21]. In [10] a new solution was suggested for homogeneous flows, of applying time-reparametrizing. This allows perfect flow estimations, in certain cases, yielding excellent linearization of the flow. In this work we follow and extend this direction. Understanding

---

<sup>\*</sup> We acknowledge support by grant agreement No. 777826 (NoMADS), by the Israel Science Foundation (Grant No. 534/19) and by the Ollendorff Minerva Center.

the time and manner of transitions of the subgradient are essential for the analysis of the TV flow. This leads both to effective mode decomposition, as well as to an alternative numerical solver. The main contributions of this work are:

1. Formulating a closed form solution of rescaled DMD (R-DMD) for TV-flow.
2. Deriving TV spectral decomposition from R-DMD.
3. Proposing a new algorithm to concisely and rapidly evaluate the TV-flow.
4. Illustrating these methods numerically, showing the alternative solution approaches yield close to identical numerical results.

## 2 Preliminary

In this section, we summarize the definitions and methods of previous studies relevant to this work.

### 2.1 *Dynamic Mode Decomposition (DMD)*

Given an observed set of  $N$  instances of length  $M$ ,  $\Psi = [\psi_0, \psi_1, \dots, \psi_N] \in \mathbb{R}^{M \times (N+1)}$ , generated by some dynamical system, the DMD algorithm [17] approximates the dynamics using a linear low-dimensional space. It finds the main spatial structures (modes), their amplitude (coefficients), and the respective time changes (eigenvalues). The three main steps of this algorithm are: 1. *Dimensionality reduction*, 2. *Optimal linear mapping*, and 3. *System reconstruction with modes, eigenvalues and coefficients*. These steps are detailed below.

**Dimensionality reduction.** It is assumed that the data is embedded in a lower dimensional space. To find this space we apply *Singular Vector Decomposition* (SVD) on the data matrix,

$$\Psi = U \Sigma V^*. \quad (1)$$

where  $V^*$  is the conjugate transpose of  $V$ .

The columns of  $U$  span the column space of  $\Psi$ , whereas the rows of  $V^*$  span the row space of  $\Psi$ . The data is projected on the subspace spanned by  $r$  columns of  $U$ , denoted  $U_r$ , related to the most significant eigenvalues,

$$x_k = U_r^* \cdot \psi_k. \quad (2)$$

Thus  $x_k$  can be understood as the coordinates of a datum  $\psi_k$  with respect to this basis.

**Linear mapping.** The linear mapping,  $F$ , from  $x_k$  to  $x_{k+1}$ , is obtained by solving the following optimization problem,  $F = \arg \min_F \sum_{k=0}^{N-1} \|F \cdot x_k - x_{k+1}\|^2$ , or, in matrix notation,

$$F = \arg \min_F \|F \cdot X - Y\|_{\mathcal{F}}^2, \quad (3)$$

where  $\|\cdot\|_{\mathcal{F}}$  denotes the Frobenius norm and

$$X = U_r^* [\psi_0 \cdots \psi_{N-1}], \quad Y = U_r^* [\psi_1 \cdots \psi_N]. \quad (4)$$

The solution to Eq. (3) is

$$F = YX^* \cdot (XX^*)^{-1}. \quad (5)$$

Thus, the relation between two successive data coordinates is given by  $x_{k+1} \approx F \cdot x_k$ , where  $\approx$  is the approximation in the sense of error minimization of Eq. (3). We assume the linear mapping,  $F$ , is diagonalizable and therefore can be written as

$$F = WDW^{-1}, \quad (6)$$

where  $W$  contains the right eigenvectors of  $F$ , and  $D$  is a diagonal matrix whose entries are the eigenvalues of  $F$ .

**Modes, eigenvalues and coefficients.** Reconstructing the datum  $\psi_{k+1}$  from the corresponding coordinates  $x_{k+1}$  is formulated as  $\tilde{\psi}_{k+1} = U_r x_{k+1}$ . Substituting  $x_{k+1}$  by  $F \cdot x_k$  we get  $\tilde{\psi}_{k+1} \approx U_r \cdot F \cdot x_k$ . By plugging in the definition of  $x_k$ , Eq. (2), we have

$$\tilde{\psi}_{k+1} \approx U_r \cdot F \cdot U_r^* \psi_k = U_r \cdot F \cdot U_r^* \tilde{\psi}_k. \quad (7)$$

Notice that the equality notation is justified as  $\tilde{\psi}_k$  is the projection of  $\psi_k$  on the subspace spanned by the columns of  $U_r$ . Moreover, by generalizing this relation we can approximate the entire dynamics as

$$\tilde{\psi}_k \approx A^k \cdot \tilde{\psi}_0, \quad (8)$$

where

$$A = U_r \cdot F \cdot U_r^*. \quad (9)$$

Substituting Eqs. (6) and (9) in Eq. (8), we get

$$\tilde{\psi}_k \approx (U_r \cdot WDW^{-1} \cdot U_r^*)^k \cdot \tilde{\psi}_0 = U_r \cdot WD^k W^{-1} \cdot U_r^* \cdot \tilde{\psi}_0.$$

Now, we can define the modes,  $\{\phi_i\}_{i=1}^r$ , eigenvalues,  $\{\mu_i\}_{i=1}^r$ , and coefficients,  $\{\alpha_i\}_{i=1}^r$ , having the dynamic mode decomposition.

*Modes* are defined as  $\Phi = [\phi_1 \cdots \phi_r] = U_r W$ . Notice that  $\{\phi_i\}_{i=1}^r$  are the right eigenvectors of the matrix  $A$  and only them since the rank of  $A$  is  $r$ .

*Eigenvalues* are the diagonal entries of the matrix  $D$ . These are the eigenvalues of the matrix  $F$  and  $A$ .

*Coefficients* are defined by  $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_r] = W^{-1} U_r^* \tilde{\psi}_0$ . We can now reconstruct the approximate dynamics by,

$$\tilde{\psi}_k \approx \Phi D^k \boldsymbol{\alpha} = \sum_{i=1}^r \alpha_i \mu_i^k \phi_i. \quad (10)$$

The DMD algorithm is summarized in Algorithm 1.

## 2.2 Total Variation Spectral Decomposition

Let  $\mathcal{H}$  be a Hilbert space with an inner product,  $\langle \cdot, \cdot \rangle$ , and the corresponding induced norm,  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ .

---

**Algorithm 1** Standard DMD [17]

---

- 1: **Inputs:** Data sequence  $\{\psi_k\}_{k=0}^N$ ; reduced dimensionality  $r$ .
  - 2: Compute the *Singular Vector Decomposition* (SVD) of  $\Psi$  (see [19]) (Eq. (1)).
  - 3: Form the matrices  $X$  and  $Y$  from the coordinates of the data (Eq. (4)).
  - 4: Find the optimal linear mapping,  $F$ , between  $X$  and  $Y$  (Eq. (5)).
  - 5: Compute the modes and the coordinates as  $\Phi \triangleq U_r W$ ,  $\alpha \triangleq W^{-1} U_r^* \psi_0$ . The eigenvalues of the DMD are the eigenvalues of  $F$ ,  $\{\mu_i\}_{i=1}^r$ .
  - 6: **Outputs:**  $\{\mu_i, \phi_i, \alpha_i\}_1^r$
- 

*TV functional.* The TV functional, is defined for smooth functions as,

$$J_{TV}(\psi) = \langle |\nabla \psi|, 1 \rangle, \quad \psi \in \mathcal{H}. \quad (11)$$

Precise definitions for functions in  $BV$  and various properties of TV can be found in [8]. We denote the subdifferential of  $J_{TV}$  at  $\psi$  as  $\partial J_{TV}(\psi)$  and a subgradient as  $-P$ . They admit the following relation,

$$P(\psi) \in -\partial J_{TV}(\psi). \quad (12)$$

We assume Neumann boundary conditions. Note that the thorough analysis of the subgradient of the TV flow can be found in [1,3].

*Eigenfunctions.* The nonlinear eigenfunction,  $v$ , of  $P$  admits

$$P(v) = \lambda \cdot v, \quad (\mathbf{EF})$$

for some  $\lambda \in \mathbb{R}$ . It can be shown (see [6]), that with the above definitions of  $P$  we have  $\lambda \leq 0$ .

*The gradient descent of TV* (TV-flow) is defined by the following PDE,

$$\psi_t = P(\psi), \quad \psi(t=0) = f, \quad (\mathbf{TV-flow})$$

where  $\psi_t$  is the first order time derivative of  $\psi(t)$  and  $f$  is the initial condition.

**TV spectral framework [13].** The spectral decomposition of a signal,  $f \in \mathcal{H}$ , related to the eigenfunctions of  $P$  is based on the solution of Eq. (**TV-flow**). We list below the definitions of the transform, inverse-transform, filtering and spectrum of this framework. We simplify notations, assume all derivatives exist and that the signal has no null-space components of  $J_{TV}$ .

*The TV-transform* is defined by  $\mathcal{G}(t) = t \frac{d^2}{dt^2} \psi(t)$ , where  $\psi(t)$  is the solution of (**TV-flow**). The function  $\mathcal{G}(t)$  is the spectral component of the signal  $f(x)$  at time  $t$ .

*The inverse transform* is the reconstruction of the original signal  $f$  from the spectral components, defined by  $\hat{f} = \int_0^\infty \mathcal{G}(t) dt$ .

*The filtering* of  $f$  by the filter  $h(t)$  is  $f_h = \int_0^\infty \mathcal{G}(t) \cdot h(t) dt$ , where  $h(t)$  is a real function. Namely, filtering is an amplification (or attenuation) of  $\mathcal{G}(t)$  in the transform domain,  $t$ .

The spectrum of  $f$  at any scale  $t$  is defined by  $S(t) = \langle f, \mathcal{G}(t) \rangle$ .

If the initial condition,  $f$ , is an eigenfunction (admits Eq. **(EF)**), then the solution of Eq. **(TV-flow)** is,

$$\psi(t) = (1 + \lambda t)^+ \cdot f, \quad (13)$$

where  $\lambda$  is the corresponding eigenvalue and  $(a)^+ = \max\{a, 0\}$ ,  $\forall a \in \mathbb{R}$ . The transform of this signal is

$$\mathcal{G}(t) = f \cdot t\lambda^2 \cdot \delta(1 + \lambda \cdot t), \quad (14)$$

where  $\delta(\cdot)$  is the Dirac delta function.

**Settings:** In this work we first note that DMD is fully discrete (time and space) whereas TV-flow and spectral TV are semi-discrete (time-continuous, spatially discrete). Thus, in order to apply DMD on a gradient descent flow we first need to sample (uniformly) with respect to the time variable  $t$ . In all cases we use Euclidean inner product and norm. We list below some *Attributes* of the semi discrete one-dimensional TV-flow and the TV spectral components:

1. The subgradient is piecewise constant with respect to  $t$  [6].
2. The initial condition can be reconstructed by knowing the subgradient as a function of  $t$  (by integration).
3. The flow splits into merging events [4].
4. The average of the subgradient,  $-P$ , over the spatial variable is zero.
5. The spectrum is a finite set of delta functions, where each delta function represents a spectral component [6].
6. For a given  $f$ , the spectral component set is orthogonal [6].
7. Two adjacent points which become equal in value during the flow, will not separate [18,3].
8. There is a time reparametrization for which the TV decays exponentially [10].

### 3 DMD of the TV flow

#### 3.1 Closed form solution

Let us formulate *Attribute 1* and *Attribute 2* more formally. The solution of **(TV-flow)** converges to a steady state in finite time. In this finite time, the solution is divided into  $L$  disjoint segments,  $[T_i, T_{i+1})$ . In each segment, the subgradient is constant,  $-p_i \in \partial J(t)$ ,  $t \in [T_i, T_{i+1})$ , where for  $t > T_L$  it is zero,  $p_{L+1} = 0$ . The solution can be expressed by (e.g. [6]),

$$\psi(t) = \psi(T_i) + (t - T_i)p_i, \quad t \in [T_i, T_{i+1}). \quad (15)$$

For an initial condition,  $\psi(0) = f$ , orthogonal to the kernel of  $J_{TV}$  (constant functions) the reconstruction of  $f$  from the set of subgradients  $\{p_i\}$  is

$$f = \psi(0) = \sum_{i=1}^L T_i(p_{i+1} - p_i). \quad (16)$$

**Proposition 1 (Linear decay).** *The solution of (TV-flow) is a sum of spectral components decaying linearly. More formally, if the initial condition,  $f$ , is orthogonal to the kernel set of  $P$  then the solution of Eq. (TV-flow) is*

$$\psi(t) = \sum_{i=1}^L (1 + \lambda_i t)^+ \varphi_i, \text{ where } \lambda_i = -T_i^{-1} \text{ and } \varphi_i = \frac{p_i - p_{i+1}}{\lambda_i}. \quad (17)$$

*Proof.* Let us reformulate the solution, Eq. (15) for the first time range  $t \in [0, T_1]$ . Substituting Eq. (16) in Eq. (15), we have

$$\begin{aligned} \psi(t) &= \psi(0) + (t-0)p_1 = \sum_{i=1}^L T_i(p_{i+1} - p_i) + tp_1 \\ &\stackrel{p_{L+1}=0}{=} \underbrace{\sum_{i=1}^L T_i(p_{i+1} - p_i)}_{\substack{= \\ \sum_{i=1}^L T_i(p_{i+1} - p_i)}} + t \sum_{i=1}^L (p_i - p_{i+1}) = \sum_{i=1}^L (T_i - t)(p_{i+1} - p_i). \end{aligned}$$

Therefore,  $\psi(T_1) = \sum_{i=1}^L (T_i - T_1)(p_{i+1} - p_i) = \sum_{i=2}^L (T_i - T_1)(p_{i+1} - p_i)$ . In a similar manner, we can reformulate the solution for  $t \in [T_1, T_2]$  as  $\psi(t)|_{t \in [T_1, T_2]} = \sum_{i=2}^L (T_i - t)(p_{i+1} - p_i)$ . By induction, the general solution is,  $\psi(t) = \sum_{i=1}^L (T_i - t)^+(p_{i+1} - p_i)$ . Denoting  $\lambda_i = -T_i^{-1}$ , this can be expressed as,

$$\psi(t) = \sum_{i=1}^L (-\lambda_i^{-1} - t)^+(p_{i+1} - p_i) = \sum_{i=1}^L (1 + \lambda_i t)^+ \frac{p_i - p_{i+1}}{\lambda_i}. \quad \square$$

The spectral decomposition is computed by second order time derivative of  $\psi(t)$ ,

$$\mathcal{G}(t) = \sum_{i=1}^L \varphi_i \cdot t \lambda_i^2 \cdot \delta(1 + \lambda_i \cdot t). \quad (18)$$

This coincides with *Attribute 5: a finite set of Dirac delta functions*. The flow can also be defined (without using the operator  $(\cdot)^+$ ) in disjoint time intervals,

$$\psi(t) = \sum_{i=k}^L (1 + \lambda_i t) \varphi_i, \quad \forall t \in [T_{k-1}, T_k]. \quad (19)$$

We will use this formulation later in our analysis.

### 3.2 Flow transitions and a fast TV-flow algorithm

While there has been on going research on fast methods for TV regularization (e.g. [11,15,9]), few advances were made in fast algorithms of the TV-flow, which is required for computing spectral TV. Our proposed solution,  $\psi(t) \in \mathbb{R}^M \times [0, T_L]$ , is in a semi-discrete setting. The algorithm is based on the TV-flow attributes listed at the end of Section 2.2.

**Algorithm 2** Accelerated TV flow and spectral decomposition (1D)

- 
- 1: **Inputs:**  $f, \delta > 0$
  - 2: **Initialize:**  $\psi_0 \leftarrow f, p_0 \leftarrow P(f), t \leftarrow 0, \mathcal{T} = \emptyset$ , and  $\mathcal{P} = \{p_0\}$
  - 3: **while**  $\|p_i\| > \delta$  **do**
  - 4: Find the next time transition point,  $T_{i+1}$  (Eq. (20)).
  - 5:  $\psi_{i+1} \leftarrow \psi_i + (T_{i+1} - T_i) \cdot p_i$ .
  - 6: Find clusters,  $\{\mathcal{M}_k\}_{k=1}^r$ , where  $|\nabla\psi_{i+1}| = 0$
  - 7: Update the next negative subgradient,  $p_{i+1}$ , such that  $p_{i+1}(\mathcal{M}_k) = \frac{1}{|\mathcal{M}_k|} \sum_{\mathcal{M}_k} p_i(\mathcal{M}_k)$ ,  $k = 1, \dots, r$ .
  - 8: Add  $p_{i+1}$  to the set  $\mathcal{P}$ ; Add  $T_{i+1}$  to the set  $\mathcal{T}$ . Compute  $\varphi_i$  by (17).
  - 9: **end while**
  - 10: **Outputs:**  $\mathcal{T}, \mathcal{P}$
- 

Subgradient transitions are time points where the subgradient is updated, denoted  $\mathcal{T} = \{T_i\}_{i=0}^L$  where  $T_0 = 0$ , and  $T_L$  is the extinction time. From *Attributes* 3 and 7 we can find these time points and compute the updated subgradient. These steps are detailed below and concisely formalized in Algorithm 2.

**Merging** occurs when two adjacent pixels become equal. Then, the discrete gradient approximation is  $\nabla\psi^{(j)} := \psi^{(j+1)} - \psi^{(j)} = 0$ , where  $\psi^{(j)}$  is the entry  $j$  of the vector  $\psi$ . Using Eq. (15), the first merging event after  $T_i$  can be calculated by

$$T_{i+1} = T_i + \min_{j \in \mathcal{J}^*} \{-\nabla\psi^{(j)}(T_i) / \nabla P^{(j)}(\psi(T_i))\}, \quad (20)$$

where  $\mathcal{J}^* = \{j \text{ s.t. } 0 < -\nabla\psi^{(j)}(T_i) / \nabla P^{(j)}(\psi(T_i)) < \infty\}$ .

**Subgradient update** is necessary after every merging event. According to *Attribute* 7, the merged entries evolve together at the same pace. In addition, the subgradient at other locations is unchanged (see [18]). Since the average of the subgradient is zero (*Attribute* 4), the subgradient of the merged entries is the average of the previous subgradient at these entries.

### 3.3 Rescaled-DMD

We follow the work of [10] where an analysis of DMD was carried out for flows based on homogeneous operators. The homogeneity order dictates not only the decay profile but also the support in time of the solution. In particular, TV-flow decays linearly and has a finite extinction time, whereas a flow linearization algorithm, such as DMD, can be interpreted as an exponential data fitting algorithm [2]. In [10] it was suggested to solve this problem by time reparametrization. Introducing a new time variable  $\tau$ , Eq. (**TV-flow**) is time rescaled by the flow,

$$\psi_\tau = G(\psi) = -\frac{\langle P(\psi), \psi \rangle}{\|P(\psi)\|^2} P(\psi). \quad (\mathbf{R-TV-flow})$$

Note that,  $G(a\psi) = aG(\psi)$ ,  $\forall a \in \mathbb{R}$ . Therefore, this can be viewed as a flow derived by a one-homogeneous operator, yielding exponential decay. Using (**TV-flow**)

and (**R-TV-flow**), the relation between  $t$  and  $\tau$  can be derived by,

$$\frac{d}{d\tau}\psi(t(\tau)) = -\frac{\langle P(\psi(t(\tau))), \psi(t(\tau)) \rangle}{\|P(\psi(t(\tau)))\|^2} P(\psi(t(\tau))) = -\frac{\langle P(\psi(t(\tau))), \psi(t(\tau)) \rangle}{\|P(\psi(t(\tau)))\|^2} \frac{d}{dt}\psi(t(\tau)),$$

yielding,

$$\frac{d}{d\tau}t(\tau) = -\frac{\langle P(\psi(t(\tau))), \psi(t(\tau)) \rangle}{\|P(\psi(t(\tau)))\|^2}. \quad (21)$$

This ODE gets a different form in each segment,  $[T_{k-1}, T_k)$ . Substituting Eq. (19) in Eq. (21), we have

$$\frac{d}{d\tau}t(\tau) = -\frac{\langle \sum_{i=k}^L \lambda_i \varphi_i, \sum_{i=k}^L (1 + \lambda_i t(\tau)) \varphi_i \rangle}{\left\| \sum_{i=k}^L \lambda_i \varphi_i \right\|^2} = -\frac{\sum_{i=k}^L \lambda_i \|\varphi_i\|^2}{\sum_{i=k}^L \lambda_i^2 \|\varphi_i\|^2} - t(\tau).$$

The solution is,

$$t(\tau) = a_k e^{-\tau} - c_k, \quad c_k = \frac{\sum_{i=k}^L \lambda_i \|\varphi_i\|^2}{\sum_{i=k}^L \lambda_i^2 \|\varphi_i\|^2}, \quad (22)$$

where  $a_k$  depends on the initial conditions of every segment such that  $t(\tau)$  is continuous (where  $t(0) = 0$ ). Then, the time points  $\{T_i\}_{i=1}^L$  are mapped to  $\{\tau_i\}_{i=1}^L$ , accordingly.

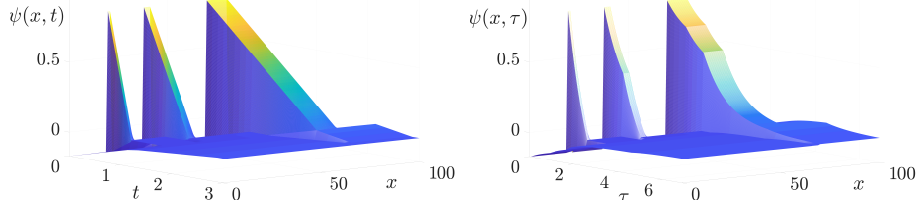


Fig. 1: **Time reparametrization** - Left, **TV-flow**  $\psi(t)$ . The TV-flow decays piecewise linearly. With time reparametrization, **R-TV-flow**,  $\psi(\tau)$ , it is mapped to a piecewise smooth function, depicted in Right. The nonsmooth points represent transitions in the subgradient.

**Proposition 2 (Main TV-flow modes).** *In every disjoint  $k$ th interval,  $[\tau_{k-1}, \tau_k)$ , the solution of time reparametrizing (**TV-flow**), Eq. (**R-TV-flow**), has two main orthogonal modes,  $\xi_1^k, \xi_2^k$ , with eigenvalues zero and minus one.*



*Proof.* Substituting Eq. (22) in Eq. (19), we get

$$\begin{aligned}
 \psi(t(\tau)) &= \sum_{i=k}^L (1 + \lambda_i t(\tau)) \varphi_i, & \forall t \in [T_{k-1}, T_k) \\
 &= \sum_{i=k}^L (1 + \lambda_i (a_k e^{-\tau} - c_k)) \varphi_i, & \forall \tau \in [\tau_{k-1}, \tau_k) \\
 &= \underbrace{\sum_{i=k}^L \varphi_i - c_k \sum_{i=k}^L \lambda_i \varphi_i}_{\xi_1^k} + e^{-\tau} \underbrace{a_k \sum_{i=k}^L \lambda_i \varphi_i}_{\xi_2^k} = \xi_1^k + e^{-\tau} \xi_2^k, \forall \tau \in [\tau_{k-1}, \tau_k).
 \end{aligned}$$

By plugging  $c_k$  from Eq. (22) into  $\xi_1^k, \xi_2^k$  their orthogonality is concluded immediately.  $\square$

In Fig. 2 we show the TV-modes defined in Prop. 2 with the initial condition Fig. 3a. It contains six disjoint intervals with the corresponding modes  $\{\xi_1^k, \xi_2^k\}_{k=1}^6$ .

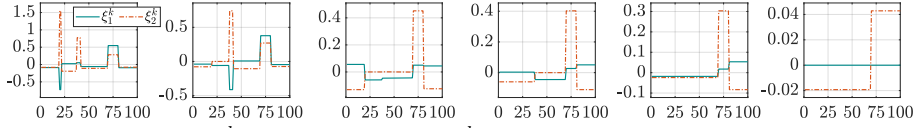


Fig. 2: Modes:  $\xi_1^k$  (teal) - constant,  $\xi_2^k$  (orange) - exponentially decaying.

### 3.4 Analysis of the Rescaled-DMD

Here, we show a closed form solution to the time Rescaled-DMD (R-DMD). The common thread in the following discussion is *Attribute 6*, the orthogonality of the TV-spectral components. The method is summarized in Algorithm 3.

**Theorem 1 (R-DMD of TV-flow).** *Let  $\tau_0 = 0$ , then for the interval,  $[\tau_{k-1}, \tau_k)$ , where  $k = 1, \dots, L-1$ , R-DMD reveals two non-zero orthogonal modes that reconstruct accurately the TV-flow in this interval. For the last interval,  $[\tau_{L-1}, \tau_L)$ , there is only one nonzero mode.*

*Proof.* According to Prop. 2 and since DMD is an exponential data fitting algorithm, the DMD of the dynamics, Eq. (R-TV-flow), is as follows. The modes are  $\phi_1^k = \xi_1^k / \|\xi_1^k\|$ ,  $\phi_2^k = \xi_2^k / \|\xi_2^k\|$ , and the coefficients are  $\alpha_1^k = \|\xi_1^k\|$  and  $\alpha_2^k = \|\xi_2^k\|$ . Note that one mode is constant with respect to time and the second decays exponentially. Therefore, the eigenvalues are  $\mu_1^k = 1$  for the constant mode and  $\mu_2^k = e^{-dt}$  where  $dt$  is the sampling step size (see Algo. 3).  $\square$

Now we formulate the relation between the TV spectral components  $\varphi_k$  and the R-DMD modes.

**Proposition 3 (Revealing TV spectral components from R-DMD).** *The  $k$ th spectral component,  $\varphi_k$ , admits the following relation*

$$c_k \lambda_k \varphi_k = \phi_2^k - \frac{\langle \phi_2^k, \phi_2^{k+1} \rangle}{\|\phi_2^{k+1}\|^2} \phi_2^{k+1}. \quad (23)$$

**Algorithm 3** R-DMD for TV-flow

- 
- 1: **Inputs:** The initial condition, and sampling step size  $dt$ .
  - 2: **Initialize:** Evolve the solution of (**R-TV-flow**) uniformly with a step size  $dt$ .
  - 3: Invoke Algo. 2 - the result is  $\mathcal{T}$  and  $\mathcal{P}$ .
  - 4: Map the set of transition time points,  $\mathcal{T}$ , to a new set  $\hat{\mathcal{T}}$  (Eq. (21)).
  - 5: **for** Every time segment  $[\tau_i, \tau_{i+1})$ ,  $\tau_i, \tau_{i+1} \in \hat{\mathcal{T}}$  **do**
  - 6:     Invoke Algo. 1 with  $r = 2$  (when  $i = L - 1$ ,  $r = 1$ ).
  - 7: **end for**
  - 8: **Outputs:** Modes  $\{\phi_1^k, \phi_2^k\}_{k=1}^L$ , coefficients  $\{\alpha_1^k, \alpha_2^k\}_{k=1}^L$ , and eigenvalues  $\mu_1^k = 1, \mu_2^k = e^{-dt}$ .
- 

*Proof.*

$$\begin{aligned}
\phi_2^k - \frac{\langle \phi_2^k, \phi_2^{k+1} \rangle}{\|\phi_2^{k+1}\|^2} \phi_2^{k+1} &= c_k \sum_{i=k}^L \lambda_i \varphi_i - \frac{\langle c_k \sum_{i=k}^L \lambda_i \varphi_i, c_{k+1} \sum_{i=k+1}^L \lambda_i \varphi_i \rangle}{\|c_{k+1} \sum_{i=k+1}^L \lambda_i \varphi_i\|^2} c_{k+1} \sum_{i=k+1}^L \lambda_i \varphi_i \\
&= c_k \lambda_k \varphi_k + c_k \sum_{i=k+1}^L \lambda_i \varphi_i - c_k \frac{\langle \lambda_k \varphi_k + \sum_{i=k+1}^L \lambda_i \varphi_i, \sum_{i=k+1}^L \lambda_i \varphi_i \rangle}{\|\sum_{i=k+1}^L \lambda_i \varphi_i\|^2} \sum_{i=k+1}^L \lambda_i \varphi_i = c_k \lambda_k \varphi_k. \quad \square
\end{aligned}$$

## 4 Results and conclusion

In this section, the theory and algorithms discussed above are illustrated. We use standard first-order discretization of the derivatives and Neumann boundary conditions. We begin with a toy example, depicted in Fig. 3a. In Fig. 1 we show

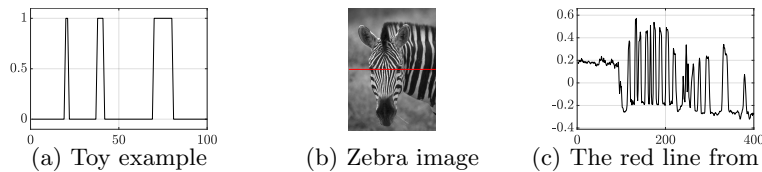


Fig. 3: **Initial conditions** - (a) Three pulses with different widths. (c) The corresponding values of the pixels on the red line in (b).

that the solution of (**TV-flow**) decays linearly and that of Eq. (**R-TV-flow**) piecewise exponentially. In Fig. 4-top the TV-spectral decomposition, dashed red line (computed in the standard way, see [13]) is compared with two algorithms. First, by Algorithm 3 based on Prop. 3, black dotted line. Second, by Algorithm 2, blue line. The errors between the TV-spectral decomposition and Algorithms 3 and 2 are depicted in Fig. 4-bottom.

In Fig. 5 we show results of the fast TV-spectral decomposition, Algorithm 2, applied on a natural signal. We arbitrarily chose the red line from the zebra in Fig. 3b, depicted in Fig. 3c. Bands of standard TV-spectral decomposition and

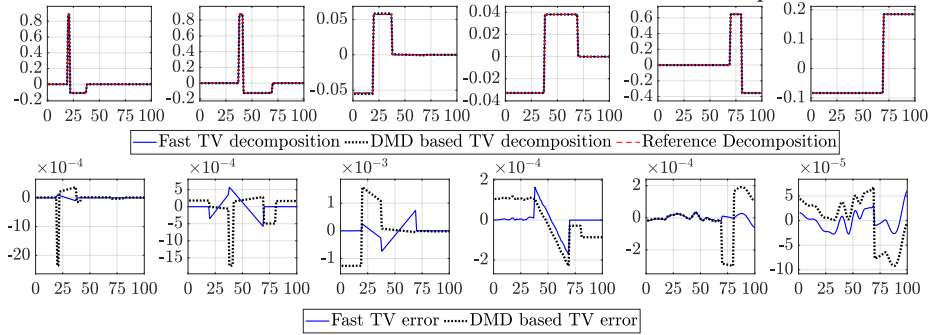


Fig. 4: **TV spectral decomposition comparison for toy example** - The standard method of spectral decomposition (Dashed red line) vs. fast TV decomposition and R-DMD decomposition in blue and dotted black lines respectively. Bottom row - respective errors

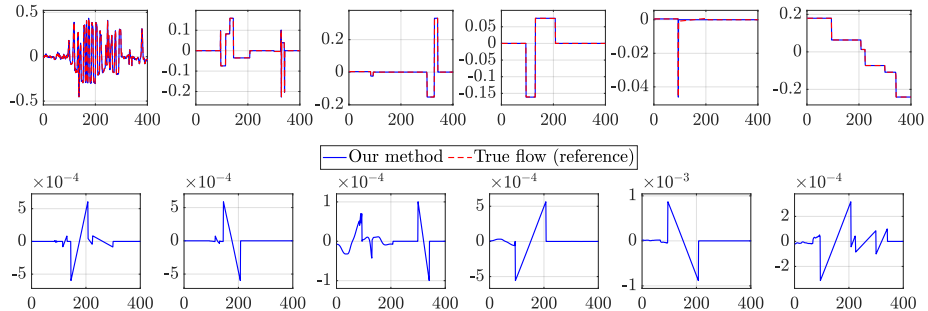


Fig. 5: **TV spectral decomposition comparison for initial condition depicted in 3c** - Standard method of spectral decomposition (Dashed red line) vs. fast TV decomposition (Blue line). Bottom row - respective error.

the fast TV-spectral decomposition are shown in Fig. 5-top. One can observe that our proposed fast method recovers the spectral bands faithfully, with negligible error, Fig. 5-bottom.

**Rough performance comparison.** We report the elapsed time in seconds, running in Matlab 2018b on an 8th Gen. Core i7 laptop with 16GB RAM. Initial condition Fig. 3a: Standard method (iterative application of [7]) - 488.4s; ours - 0.013s. Initial condition Fig. 3c (zebra): Standard method -  $7.1 \times 10^3$ s; ours - 0.15s. Note that the times specified for our algorithm do not include the initialization of  $p_0$ , which takes additional 7.3s for Fig. 3a, and 27.8s for Fig. 3c. This time can be shortened further using faster subgradient computation algorithms.

**Conclusion.** In this paper the modes of one-dimensional TV-flow were analyzed. A popular method for finding modes of flows in fluid dynamics, *Dynamic Mode Decomposition* (DMD) [17], was examined. We have presented an adaptation for the TV case, by a time rescaled version. This was based on the spectral-TV theory, where the spectral components are orthogonal. Obtaining TV modes, or nonlinear spectral TV components, requires a gradient-flow evolu-

tion. Since evolving TV-flow is a slow process using optimization techniques, we have proposed a very fast method, based on simple updates of the subgradient.

## References

1. Andreu, F., Ballester, C., Caselles, V., Mazón, J.M.: Minimizing total variation flow. *Differential and Integral Equations* **14**(3), 321–360 (2001)
2. Askham, T., Kutz, J.N.: Variable projection methods for an optimized dynamic mode decomposition. *SIAM Journal on Appl. Dyn. Sys.* **17**(1), 380–416 (2018)
3. Bellettini, G., Caselles, V., Novaga, M.: The total variation flow in rn. *Journal of Differential Equations* **184**(2), 475–525 (2002)
4. Brox, T., Weickert, J.: A TV flow based local scale estimate and its application to texture discrimination. *J. of Vis. Comm. and Image Rep.* **17**(5), 1053–1073 (2006)
5. Bungert, L., Burger, M., Chambolle, A., Novaga, M.: Nonlinear spectral decompositions by gradient flows of one-homogeneous functionals. *arXiv preprint arXiv:1901.06979* (2019)
6. Burger, M., Gilboa, G., Moeller, M., Eckardt, L., Cremers, D.: Spectral decompositions using one-homogeneous functionals. *SIAM Im. Sci.* **9**(3), 1374–1408 (2016)
7. Chambolle, A.: An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision* **20**(1-2), 89–97 (2004)
8. Chambolle, A., Caselles, V., Cremers, D., Novaga, M., Pock, T.: An introduction to total variation for image analysis. *Theoretical foundations and numerical methods for sparse recovery* **9**(263-340), 227 (2010)
9. Cherkaoui, H., Sulam, J., Moreau, T.: Learning to solve tv regularised problems with unrolled algorithms. *Adv.Neural Inf. Proc. Sys.* **33** (2020)
10. Cohen, I., Azencot, O., Lifshits, P., Gilboa, G.: Modes of homogeneous gradient flows. *arXiv preprint arXiv:2007.01534* (2020)
11. Darbon, J., Sigelle, M.: Image restoration with discrete constrained total variation part i: Fast and exact optimization. *J. Math. Im. and Vision* **26**(3), 261–276 (2006)
12. Gilboa, G.: A spectral approach to total variation. In: *Inter. Conf. on Scale Space and Variational Methods in Computer Vision*. pp. 36–47. Springer (2013)
13. Gilboa, G.: A total variation spectral framework for scale and texture analysis. *SIAM journal on Imaging Sciences* **7**(4), 1937–1961 (2014)
14. Gilboa, G., Moeller, M., Burger, M.: Nonlinear spectral analysis via one-homogeneous functionals: Overview and future prospects. *J. of Math. Im. and Vision* **56**(2), 300–319 (2016)
15. Goldfarb, D., Yin, W.: Parametric maximum flow algorithms for fast total variation minimization. *SIAM J. Sci. Comp.* **31**(5), 3712–3743 (2009)
16. Mezić, I.: Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics* **41**(1-3), 309–325 (2005)
17. Schmid, P.J.: Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics* **656**, 5–28 (2010)
18. Steidl, G., Weickert, J., Brox, T., Mrázek, P., Welk, M.: On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and sides. *SIAM J. on Num. Ana.* **42**(2), 686–713 (2004)
19. Trefethen, L.N., Bau III, D.: *Numerical linear algebra*, vol. 50. Siam (1997)
20. Williams, M.O., Kevrekidis, I.G., Rowley, C.W.: A data-driven approximation of the koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science* **25**(6), 1307–1346 (2015)
21. Williams, M.O., Rowley, C.W., Kevrekidis, Y.: A kernel-based method for data-driven koopman spectral analysis. *J. of Comp. Dyn.* **2**(2), 247–265 (2015)