J Math Chem (2015) 53:1885–1914
DOI 10.1007/s10910-015-0527-1

ORIGINAL PAPER

Sparse and nonnegative sparse D-MORPH regression

Genyuan Li[1] · **Roberto Rey-de-Castro**[1] · **Xi Xing**[1] · **Herschel Rabitz**[1]

Received: 21 April 2015 / Accepted: 24 June 2015 / Published online: 4 July 2015

Abstract An underdetermined linear algebraic equation system $\mathbf{y} = \mathbf{\Phi x}$, where $\mathbf{\Phi}$ is an $m \times n (m < n)$ rectangular constant matrix with rank $r \leq m$ and $\mathbf{y} \in \text{Ran}(\mathbf{\Phi})$ (range of $\mathbf{\Phi}$), has an infinite number of solutions. Diffeomorphic modulation under observable response preserving homotopy (D-MORPH) regression seeks a solution satisfying the extra requirement of minimizing a chosen cost function, $\mathcal{K}$. A wide variety of choices of the cost function makes it possible to achieve diverse goals, and hence D-MORPH regression has been successfully applied to solve a range of problems. In this paper, D-MORPH regression is extended to determine a *sparse* or a *nonnegative sparse* solution of the vector $\mathbf{x}$. For this purpose, recursive reweighted least-squares (RRLS) minimization is adopted and modified to construct the cost function $\mathcal{K}$ for D-MORPH regression. The advantage of sparse and nonnegative sparse D-MORPH regression is that the matrix $\mathbf{\Phi}$ does not need to have row-full rank, thereby enabling flexibility to search for sparse solutions $\mathbf{x}$ with ancillary properties in practical applications. These tools are applied to (a) simulation data for quantum-control-mechanism identification utilizing high dimensional model representation (HDMR) modeling and (b) experimental mass spectral data for determining the composition of an unknown mixture of chemical species.

✉ Herschel Rabitz
hrabitz@princeton.edu

Genyuan Li
genyuan@princeton.edu

Roberto Rey-de-Castro
rrey@princeton.edu

Xi Xing
xxing@princeton.edu

[1] Department of Chemistry, Princeton University, Princeton, NJ 08544, USA

Keywords Underdetermined system · D-MORPH regression · Least-squares regression · IRLS · RRLS · Quantum-control-mechanism identification · Mass spectrum analysis

1 Introduction

The linear algebraic equation system

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x}, \tag{1}$$

where $\mathbf{\Phi}$ is an $m \times n (m < n)$ rectangular constant matrix with rank $r (r \leq m)$, and $\mathbf{y} \in \mathrm{Ran}(\mathbf{\Phi})$ (range of $\mathbf{\Phi}$), is underdetermined and has an infinite number of solutions $\mathbf{x}$ composing an $(n - r)$-dimensional completely connected manifold $\mathcal{M}$. "In the absence of any other information, any solution of Eq. 1 has an equal preference over all others. However, many scientific applications work under the condition that the desired solution $\mathbf{x} \in \mathcal{M}$ is either sparse or well approximated by a sparse vector(s)" [1]. Hereafter, we denote that a vector $\mathbf{x}$ has sparsity k (or is k-sparse) if it has *at most* k nonzero elements. Since $\mathbf{\Phi}$ has rank r, Eq. 1 may be r-sparse. It can be proved that the sparse solutions $\mathbf{x} \in \mathcal{M}$ of Eq. 1 are those with minimal ℓ_0 or ℓ_1 norm [1–5]:

$$\mathbf{x} = \underset{\mathbf{x}\in\mathcal{M}}{\arg\min} \ \|\mathbf{x}\|_{\ell_p}, \tag{2}$$

where $p = 0, 1$. Here the ℓ_p norm is defined as

$$\|\mathbf{x}\|_{\ell_p} = \begin{cases} \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, & 0 \leq p < \infty, \\ \max\limits_{i=1,\ldots,n} |x_i|, & p = \infty. \end{cases} \tag{3}$$

Solving underdetermined systems by ℓ_1-minimization is the heart of many numerical algorithms for approximation, compression, and statistical estimation when there is a k-sparse solution for Eq. 1. Under certain assumptions on $\mathbf{\Phi}$ and $\mathbf{y}$, Eq. 2 has a unique solution $\mathbf{x}^*$, which may be found by linear programming [1,6,7]. More efficient and simpler algorithms than standard linear programming have also been considered [8,9].

Daubechies et al. proposed an alternative method of determining $\mathbf{x}^*$. They proved that if Eq. 2 has a solution $\mathbf{x}^*$ with minimal ℓ_1 norm and no vanishing elements, then the unique solution $\mathbf{x}^{\mathbf{w}}$ of the weighted least-squares problem

$$\mathbf{x}^{\mathbf{w}} = \underset{\mathbf{x}\in\mathcal{M}}{\arg\min} \ \mathcal{J}(\mathbf{w}, \mathbf{x}) = \underset{\mathbf{x}\in\mathcal{M}}{\arg\min} \ \frac{1}{2}\sum_{i=1}^{n} w_i x_i^2, \tag{4}$$

where

$$\mathbf{w} = (w_1, w_2, \ldots, w_n), \qquad w_i = |x_i^*|^{-1}$$

given by

$$\mathbf{x}^{\mathbf{w}} = W^{-1}\mathbf{\Phi}^T(\mathbf{\Phi}W^{-1}\mathbf{\Phi}^T)^{-1}\mathbf{y} \tag{5}$$

coincides with $\mathbf{x}^*$ when $\mathbf{\Phi}$ satisfies the restricted isometry property (RIP) [1]. Here, $W = \mathrm{diag}(w_1, w_2, \ldots, w_n)$. Since $\mathbf{x}^*$ is unknown, this observation cannot be directly used. However, it leads to the following paradigm for finding $\mathbf{x}^*$: (1) choose a starting weight $\mathbf{w}_0$ and solve Eq. 4 with this weight, (2) use the solution $\mathbf{x}^{(1)}$ to define a new weight $\mathbf{w}_1$ and repeat this process. This method is referred as *iteratively reweighted least-squares* (IRLS) minimization. Daubechies et al. [1] proved that when the solution $\mathbf{x}^{(n)}$ at the $(n-1)$th step is sufficiently close to the limit, the remaining steps of the algorithm converge exponentially fast (i.e., linear convergence in the terminology of numerical optimization).

Yagle [2] proposed a variant of IRLS to minimize

$$\|\mathbf{x}\|_p^p = \sum_{i=1}^{n} \frac{1}{|x_i|^{2-p}} x_i^2 \tag{6}$$

recursively, not iteratively, by slightly reducing the norm order p by $\delta (0 < \delta < 1)$ at each recursion. At step $l (0, 1, \ldots, 2/\delta)$, the weight $\mathbf{w}$ is set to be

$$w_i = \frac{1}{|x_i^{(l)}|^{l\delta} + \varepsilon}, \qquad i = 1, 2, \ldots, n, \tag{7}$$

where $\mathbf{x}^{(l)}$ is the solution of weighted least-squares regression at step $l-1$, and ε is a small number (e.g., 10^{-5}, 10^{-7}) to avoid a singularity when $x_i^{(l)} = 0$. Initially ($l = 0$) the objective function

$$\mathcal{J}(\mathbf{w}, \mathbf{x}) = \frac{1}{2}\sum_{i=1}^{n} w_i x_i^2 \approx \frac{1}{2}\sum_{i=1}^{n} x_i^2 \tag{8}$$

corresponds to minimizing the ℓ_2 norm of $\mathbf{x}$ giving the least-squares regression solution $\mathbf{x}^{(1)}$ of Eq. 1. Substituting the resultant vector $\mathbf{x}^{(1)}$ into Eq. 7 to construct new w_i's, then $\mathbf{x}^{(2)}$ is obtained from Eq. 5. The treatment is repeated until $l = 2/\delta$. The objective function at the final step

$$\mathcal{J}(\mathbf{w}, \mathbf{x}) = \frac{1}{2}\sum_{i=1}^{n} \frac{x_i^2}{|x_i^{(2/\delta)}|^2 + \varepsilon} \approx \frac{1}{2}\sum_{i=1}^{n} |x_i|^0 \tag{9}$$

equals one half of the number of nonzero x_i's. Minimizing $\mathcal{J}(\mathbf{w}, \mathbf{x})$ implies that the number of nonzero x_i's is minimal, i.e., the sparsest solution of $\mathbf{x}$. This result is valid if $\mathbf{x}^{(2/\delta)}$ converges to the sparsest solution $\mathbf{x}$. If convergence is not satisfactory, one can reduce δ. No strict proof of convergence for this recursive reweighted least-squares

(RRLS) procedure has been given. Nevertheless, RRLS minimization often converges in practice. The advantage of RRLS minimization is that only a few recursions are typically needed. For example, if setting $\delta = 0.2$, only ten recursions are performed. Furthermore, making some modification (shown in the next section) in the definition of $\mathcal{J}(\mathbf{w}, \mathbf{x})$, the solution $\mathbf{x}$ is not only sparse, but also only has nonnegative elements, which is relevant in some applications.

Both IRLS and RRLS minimizations used Eq. 5 to find the weighted least-squares solution $\mathbf{x}^{\mathbf{w}}$, which requires that $\mathbf{\Phi}$ is of row-full rank. Even if $\mathbf{\Phi}$ satisfies this condition, the matrix $\mathbf{\Phi} W^{-1} \mathbf{\Phi}^T$ can be close to singular, and evaluation of the inverse $(\mathbf{\Phi} W^{-1} \mathbf{\Phi}^T)^{-1}$ may cause a large error. This shortcoming can be avoided by using D-MORPH regression which uses singular value decomposition without the requirement of $\mathbf{\Phi}$ having row-full rank. This is beneficial in practical applications where the row-full rank of $\mathbf{\Phi}$ may not be guaranteed. Thus, in the present paper we propose a new method which combines D-MORPH regression and RRLS minimization. The weight used in RRLS minimization is adopted and modified in the construction of the cost function $\mathcal{K}$ for D-MORPH regression to find either a sparse or nonnegative sparse solution.

High dimensional model representation (HDMR) is a general set of tools to treat high dimensional input–output behavior [10–12]. As the contributions of the d-dimensional input variables $\mathbf{z}$ upon the output $f(\mathbf{z})$ can be independent and cooperative, HDMR expresses $f(\mathbf{z})$ as a finite hierarchical expansion:

$$f(\mathbf{z}) = f_0 + \sum_{\emptyset \neq u \subseteq d} f_u(\mathbf{z}_u), \tag{10}$$

where $u \subseteq \{1, 2, \ldots, d\}$ (for simplicity, we will write it as $u \subseteq d$), and $\mathbf{z}_u$ are the variables in $\mathbf{z}$ whose indexes are in u. In practice, the HDMR component functions are approximated by suitable basis functions, and $f(\mathbf{z})$ becomes a linear combinations of these basis functions

$$f(\mathbf{z}) = f_0 + \sum_{\emptyset \neq u \subseteq d} \sum_{k=1}^{k_u} c_{uk} \phi_{uk}(\mathbf{z}_u), \tag{11}$$

here c_{uk}'s are constant combination coefficients, k_u is an integer, $\phi_{uk}(\mathbf{z}_u)$'s are polynomials, splines, etc. If m points of the input–output data are sampled, i.e.,

$$f(\mathbf{z}^{(s)}) = f_0 + \sum_{\emptyset \neq u \subseteq d} \sum_{k=1}^{k_u} c_{uk} \phi_{uk}(\mathbf{z}_u^{(s)}), \qquad s = 1, 2, \ldots, m, \tag{12}$$

setting $y^{(s)} = f(\mathbf{z}^{(s)}) - f_0$ and $\mathbf{x}$ being all combination coefficients c_{uk}'s, Eq. 12 just gives Eq. 1 when m is less than the total number n of coefficients c_{uk}'s.

The c_{uk}'s can be determined by solving Eq. 1 with regression. Generally, the total number n of coefficients c_{uk}'s is large. Consider an example of $d = 10$ variables. If the 3rd order truncated HDMR expansion is used with polynomial basis functions having

degree up to 3, the total number n of coefficients c_{uk}'s is 3675. When the extended basis is used for correlated variables [12], the total number of coefficients c_{uk}'s is even larger. To accurately determine these unknown coefficients by least-squares regression, significantly more data than unknown coefficients are needed, which is often infeasible. Fortunately, D-MORPH regression can treat the case where the number of data is *less than* the number of unknowns if the cost function $\mathcal{K}$ is properly chosen [13,14]. A system with more than 3500 unknowns has been satisfactorily solved with only ∼100 data even without the requirement that the c_{uk}'s are sparse [15].

However, for many systems the HDMR expansion is intrinsically sparse. One such example is the quantum-control simulation given in Sect. 3. We will show that the approach presented here of sparse and nonnegative sparse D-MORPH regression, can be employed to find the proper solution.

The paper is organized as follows. Section 2 presents the methodology of sparse and nonnegative sparse D-MORPH regression. In Sect. 3 two examples are used to illustrate the application of sparse and nonnegative sparse D-MORPH regression: (1) the mechanism identification of quantum-control simulation data is presented to illustrate sparse D-MORPH regression; (2) chemical composition analysis of an unknown mixture utilizing mass spectral data is used to illustrate nonnegative sparse D-MORPH regression. Finally, Sect. 4 presents concluding remarks.

2 Sparse and nonnegative sparse D-MORPH regressions

The principles of D-MORPH regression are briefly summarized here; further details may be found in references [13,14]. When the number n of unknown parameters (x_i's) is larger than the rank r of $\mathbf{\Phi}$ and $\mathbf{y} \in \mathrm{Ran}(\mathbf{\Phi})$, then Eq. 1 is consistent and has an infinite number of solutions $\mathbf{x}$ with the general form

$$\mathbf{x} = \mathbf{\Phi}^{-}\mathbf{y} + (\mathrm{I}_n - \mathbf{\Phi}^{-}\mathbf{\Phi})\mathbf{u}, \tag{13}$$

where I_n is the identity matrix of dimension n and $\mathbf{u}$ is an arbitrary vector in $\mathbb{R}^n$, and $\mathbf{\Phi}^{-}$ is a generalized inverse of $\mathbf{\Phi}$ satisfying the condition

$$\mathbf{\Phi}\mathbf{\Phi}^{-}\mathbf{\Phi} = \mathbf{\Phi}. \tag{14}$$

One choice for $\mathbf{\Phi}^{-}$ in Eq. 13 is $\mathbf{\Phi}^{+}$ (which is unique) satisfying all four Penrose conditions [16]. Then,

$$\mathbf{x} = \mathbf{\Phi}^{+}\mathbf{y} + (\mathrm{I}_n - \mathbf{\Phi}^{+}\mathbf{\Phi})\mathbf{u}, \tag{15}$$

and Eq. 15 with $\mathbf{u} = \mathbf{0}$ (i.e., $\mathbf{\Phi}^{+}\mathbf{y}$) is the solution from traditional least-squares regression with the minimal norm $\|\mathbf{x}\|_{\ell_2}$.

All the solutions $\mathbf{x}$ of Eq. 1 given by Eq. 15 compose an $(n - r)$-dimensional completely connected manifold $\mathcal{M} \subset \mathbb{R}^n$. D-MORPH regression seeks a solution satisfying an extra requirement by considering an exploration path $\mathbf{x}(s)$ within $\mathcal{M}$ with $s \in [0, \infty)$, described by the differential equation

$$\frac{\mathrm{d}\mathbf{x}(s)}{\mathrm{d}s} = (\mathrm{I}_n - \mathbf{\Phi}^+\mathbf{\Phi})\frac{\mathrm{d}\mathbf{u}(s)}{\mathrm{d}s} = (\mathrm{I}_n - \mathbf{\Phi}^+\mathbf{\Phi})\mathbf{v}(s) = P\mathbf{v}(s), \tag{16}$$

where P is an orthogonal projector satisfying

$$P = P^2, \qquad P = P^T, \qquad P = PP = P^TP. \tag{17}$$

The function vector $\mathbf{v}(s)$ may be freely chosen to not only enable broad choices for exploring $\mathbf{x}(s)$, but also provide the possibility of continuously reducing a defined cost function $\mathcal{K}(\mathbf{x}(s))$ (e.g., the model variance, fitting smoothness, the weighted norm of $\mathbf{x}$, etc.) along the exploration path. This can be achieved by choosing the free function as

$$\mathbf{v}(s) = -\frac{\partial\mathcal{K}(\mathbf{x}(s))}{\partial\mathbf{x}}. \tag{18}$$

Then, we obtain

$$\frac{\mathrm{d}\mathcal{K}(\mathbf{x}(s))}{\mathrm{d}s} = \left(\frac{\partial\mathcal{K}(\mathbf{x}(s))}{\partial\mathbf{x}}\right)^T\frac{\mathrm{d}\mathbf{x}(s)}{\mathrm{d}s} = \left(\frac{\partial\mathcal{K}(\mathbf{x}(s))}{\partial\mathbf{x}}\right)^T P\mathbf{v}(s) = -\left(P\frac{\partial\mathcal{K}(\mathbf{x}(s))}{\partial\mathbf{x}}\right)^T\left(P\frac{\partial\mathcal{K}(\mathbf{w}(s))}{\partial\mathbf{x}}\right) \le 0, \tag{19}$$

i.e., the cost $\mathcal{K}$, used as an additional requirement, will be continuously reduced (systematically refining the model) over the exploration course for $s \ge 0$. Therefore,

$$\mathbf{x}_\infty = \lim_{s\to\infty}\mathbf{x}(s)$$

is the solution with a minimum value of $\mathcal{K}$. When the cost function is defined as a quadratic form in $\mathbf{x}$

$$\mathcal{K} = \frac{1}{2}\mathbf{x}^TC\mathbf{x}, \tag{20}$$

where C is symmetric and nonnegative definite, the analytical form of $\mathbf{x}_\infty$ has been obtained as [13]

$$\mathbf{x}_\infty = V_{n-r}(U_{n-r}^TV_{n-r})^{-1}U_{n-r}^T\mathbf{\Phi}^+\mathbf{y}, \tag{21}$$

where U_{n-r}, and V_{n-r} are the last $n-r$ columns of U and V obtained by singular value decomposition of PC

$$PC = U\begin{bmatrix} S_r & 0 \\ 0 & 0 \end{bmatrix}V^T \tag{22}$$

with S_r being an r-dimensional diagonal matrix of nonzero singular values.

Equation 21 is the key practical formula for the optimal solution $\mathbf{x}$ obtained by D-MORPH regression to replace Eq. 5. This solution $\mathbf{x}_\infty$ is unique in $\mathcal{M}$ corresponding to the global minimum of the cost function $\mathcal{K}$. As evident in Eq. 21, the solution $\mathbf{x}_\infty$ given by D-MORPH regression is a special linear combination of the elements of $\mathbf{x}$ obtained by least-squares regression (i.e., $\mathbf{\Phi}^+\mathbf{y}$).

For sparse and nonnegative sparse solutions of D-MORPH (referred to as sD-MORPH and nnsD-MORPH, respectively) regression, a special cost function, similar to the objective function $\mathcal{J}(\mathbf{w}, \mathbf{x})$ used in RRLS, is defined to minimize $\|\mathbf{x}\|_{\ell_0}$. When C is a diagonal matrix, the cost function in Eq. 20 becomes

$$\mathcal{K} = \frac{1}{2}\sum_{i=1}^{n} c_i x_i^2. \tag{23}$$

Similar to RRLS [2], we define

$$c_i = \frac{1}{|x_i^{(l)}|^{l\delta} + \varepsilon}, \qquad l = 0, 1, \ldots, 2/\delta, \tag{24}$$

under the condition

$$x_i^{(l)} = x_i^{(l)}, \ \forall i, \qquad \text{sD-MORPH}, \tag{25}$$

$$\begin{cases} x_i^{(l)} = x_i^{(l)}, \text{ if } \ x_i^{(l)} \geq 0, \\ x_i^{(l)} = 0, \quad \text{if } \ x_i^{(l)} < 0. \end{cases} \qquad \text{nnsD-MORPH}. \tag{26}$$

Setting $x_i^{(l)} = 0$ for negative $x_i^{(l)}$ gives it the largest shrinkage weight $1/\varepsilon$ tending to force x_i to approach zero and thereby removing negative elements in the solution $\mathbf{x}$ [17].

sD-MORPH and nnsD-MORPH regression is similar to RRLS minimization except that (1) for nnsD-MORPH regression setting $x_i^{(l)} = 0$ for negative $x_i^{(l)}$ in the weight function, (2) using Eq. 21 to replace Eq. 5 for determining the weighted least-squares solution $\mathbf{x}$. Equation 21 uses singular value decomposition instead of the inverse $(\mathbf{\Phi} W^{-1} \mathbf{\Phi}^T)^{-1}$ to avoid a large error when $\mathbf{\Phi} W^{-1} \mathbf{\Phi}^T$ is singular or near singular. Equation 21 also contains a term with the inverse $(U_{n-r}^T V_{n-r})^{-1}$, but it can be proved that $U_{n-r}^T V_{n-r}$ is nonsingular and its inverse can be accurately determined. Therefore, sD-MORPH and nnsD-MORPH regression is more flexible than IRLS and RRLS minimization in searching for sparse and nonnegative sparse solutions. This is especially beneficial in practical applications where the row-full rank of $\mathbf{\Phi}$ may not be guaranteed.

3 Illustrations of sparse and nonnegative sparse D-MORPH regression

Two examples will be used to illustrate the application of sparse and nonnegative sparse D-MORPH regression. Mechanism identification of quantum-control simulation data is presented for illustration of sparse D-MORPH regression, and chemical composition analysis of an unknown mixture utilizing mass spectral experimental data is used for illustration of nonnegative sparse D-MORPH regression.

3.1 Application to mechanism identification of quantum-control simulation data

Quantum-control seeks to manipulate dynamical events at the atomic and molecular scale using tailored laser fields [18]. An automatic closed-loop procedure is often used to find laser fields that maximize a signal reflecting the yield in a desired final state [19]. In this way quantum systems can be directed to perform a variety of tasks. In order to gain an understanding of the dynamical mechanisms induced by the fields in the quantum system under control, the technique of Hamiltonian encoding and observable decoding (HE-OD) has been implemented in the laboratory [20–22]. Over a sequence of tailored experiments, HE-OD introduces special encoded signatures into the spectral components of the control field, and the outcome appears as a modulated (encoded) signal. Decoding the modulated signal identifies the hierarchy of correlations between the components of the control field. The HE-OD procedure yields the complex amplitudes corresponding to terms in the Dyson expansion for the time evolution operation [22]. Identification of the hierarchy of correlations between components of the control field from the signals can reveal the mechanism [20]. This task can also be achieved by using sD-MORPH regression as shown below.

Here we consider a simulated system used to model atomic Rb under the excitation by a femtosecond laser pulse [22]. The simulated laser spectrum was approximately a Gaussian centered at 780 and 40 nm full width at half maximum. Under these conditions, the Rb system can be approximated by a 4-level system with states $|k\rangle$ and energies $\epsilon_k (k = 1, 2, 3, 4)$ [22]. Population can be exchanged among the various states through the absorption or emission of a photon whose frequency corresponds to the energy difference between the two states, which allows for controlling the various transitions through manipulation of the spectral phases or amplitudes of the laser field. The output signal θ was taken as the final population of state $|4\rangle$.

3.1.1 Construction of high dimensional model representation (HDMR) expansion

The phases $\phi_k (k = 1, 2, \ldots, d)$ of selected spectral components of the laser field are encoded [21,22], which yields N new laser pulse shapes that are then applied to drive the evolution of the quantum system with the outcome in state $|4\rangle$ recoded for each laser pulse. Note that the number of spectral phase components can be larger than the number of transitions in Rb, as each transition has breadth generally larger than the spectral resolution of the instrument pulse sharper. Thus, physically the effects of manipulating the selected spectral components on the output signal depend on the proximity of their corresponding photon energies to the electronic transition energies. The measurements of the resulting observed laboratory output signals $\theta(s), (s = 1, 2, \ldots, N)$ of the modulated quantum system can be expressed as [14]

$$\theta(\phi(s)) \approx |a_0 + \sum_{j=1}^{m_p} a_j e^{i(\mathbf{m}_j^T \phi(s))}|^2, \qquad (s = 1, 2, \ldots, N) \tag{27}$$

where $a_j = r_j e^{i\varphi_j}$ $(j = 0, 1, 2, \ldots, m_p)$ are complex numbers, and vectors

$$\phi(s) = (\,\phi_1(s)\ \phi_2(s)\ \cdots\ \phi_d(s)\,)^T, \tag{28}$$

$$\mathbf{m}_j = (\,m_{j1}\ m_{j2}\ \cdots\ m_{jd}\,)^T. \tag{29}$$

Here the vector $\mathbf{m}_j$ consists of (positive, zero and negative) integers $m_{jk} \in \mathbb{Z}$ representing different transition processes [20]. To be physically meaningful (i.e., some transitions actually occur) $\mathbf{m}_j$ is not a null vector (i.e., not all m_{jk} are zero). Finally, the number m_p denotes all possible transition processes.

In practice, an *a priori* physically based mechanism may not be known. Hence, we consider as many processes as possible such that Eq. 27 is big enough to likely contain the physically motivated "true" mechanism. This can be achieved by choosing a sufficiently large integer K and including all $\mathbf{m}_j$'s in Eq. 27 whose ℓ_1 norms $\|\mathbf{m}_j\|_{\ell_1} = \sum_{k=1}^{d} |m_{jk}|$ are not larger than K. The terms with large modulus of a_j in Eq. 27 identify the set of relevant quantum pathways (i.e., the mechanism, see references [20–22] for a detailed physical interpretation of such mechanisms). Thus, determination of the *complex* numbers $a_j (j = 0, 1, \ldots, m_p)$ corresponds to revealing the mechanism, which is obtained in a two step process starting from the identification of the set of the spectral correlations $\{\alpha, \beta\}$ (see Eq. 31) present in the output signal [22]. Under some conditions [22], it is possible to substitute $\{\alpha, \beta\}$ into a system of quadratic equations to reveal the mechanism. Below we use sD-MORPH regression to obtain $\{\alpha, \beta\}$ from simulated quantum-control data.

Expanding the square modulus in Eq. 27 results in the following expression [14]:

$$\begin{aligned}
\theta(\phi(s)) &\approx |r_0 e^{i\varphi_0} + \sum_{j=1}^{m_p} r_j e^{i\varphi_j} e^{i(\mathbf{m}_j^T \phi(s))}|^2 = |r_0 e^{i\varphi_0} + \sum_{j=1}^{m_p} r_j e^{i(\varphi_j + \mathbf{m}_j^T \phi(s))}|^2 \\
&= a_0 + \sum_{j=1}^{m_p} \left[a_j \cos\left(\mathbf{m}_j^T \phi(s)\right) + b_j \sin\left(\mathbf{m}_j^T \phi(s)\right) \right] \\
&\quad + \sum_{1 \le p < q \le m_p} \left[a_{pq} \cos\left((\mathbf{m}_p - \mathbf{m}_q)^T \phi(s)\right) - b_{pq} \sin\left(\mathbf{m}_p - \mathbf{m}_q)^T \phi(s)\right) \right], \\
&\quad (s = 1, 2, \ldots, N)
\end{aligned} \tag{30}$$

where $\{a, b\}$ are real constant parameters composed of the products of $r_j, \cos(\varphi_p - \varphi_q)$ and $\sin(\varphi_p - \varphi_q)$; Some $\mathbf{m}_j$ and $\mathbf{m}_p - \mathbf{m}_q$ may be equal and their corresponding sine and cosine functions will be combined together. Let the $\mathbf{n}_j$'s represent all n_p distinct vectors obtained from $\mathbf{m}_j$ and $\mathbf{m}_p - \mathbf{m}_q$. Then, Eq. 30 can be written as

$$\begin{aligned}
\theta(\phi(s)) &\approx \alpha_0 + \sum_{j=1}^{n_p} \left[\alpha_j \cos\left(\mathbf{n}_j^T \phi(s)\right) + \beta_j \sin\left(\mathbf{n}_j^T \phi(s)\right) \right], \\
&\quad (s = 1, 2, \ldots, N).
\end{aligned} \tag{31}$$

The terms with large magnitudes of α_j and β_j reveal the important physical processes $\mathbf{n}_j$ resulting from $\mathbf{m}_j$ or $\mathbf{m}_p - \mathbf{m}_q$. Equation 31 shows that the relation between θ and ϕ is an expansion of multi-variate discrete sine and cosine functions. Depending on the

number of nonzero elements in $\mathbf{n}_j$, Eq. 31 may be divided into sine and cosine functions containing $1, 2, \ldots, d$ ϕ_k's. The magnitudes of their coefficients $\{\alpha, \beta\}$ reflect the weights of the contributions of processes $\mathbf{n}_j$ with different numbers of nonzero elements. Therefore, direct comparison of the magnitudes of $\{\alpha, \beta\}$ can provide the information of the physically important transition processes.

Denoting ϕ_i as z_i, ϕ as $\mathbf{z}$ and θ as y, then Eq. 31 can be expanded and reorganized as

$$
\begin{aligned}
y(\mathbf{z}) = \alpha_0 &+ \sum_{i=1}^{d}\sum_{n_i=1}^{K}\left[\alpha_{n_i}^{(i)}\sin(n_i z_i) + \beta_{n_i}^{(i)}\cos(n_i z_i)\right] \\
&+ \sum_{1\le i<j\le d}\sum_{n_i=1}^{K}\sum_{\substack{n_j=-K\\ n_j\neq 0}}^{K}\left[\alpha_{n_i n_j}^{(ij)}\sin(n_i z_i + n_j z_j) + \beta_{n_i n_j}^{(ij)}\cos(n_i z_i + n_j z_j)\right] \\
&+ \sum_{1\le i<j<k\le d}\sum_{n_i=1}^{K}\sum_{\substack{n_j=-K\\ n_j\neq 0}}^{K}\sum_{\substack{n_k=-K\\ n_k\neq 0}}^{K}\left[\alpha_{n_i n_j n_k}^{(ijk)}\sin(n_i z_i + n_j z_j + n_k z_k)\right. \\
&\left. + \beta_{n_i n_j n_k}^{(ijk)}\cos(n_i z_i + n_j z_j + n_k z_k)\right] + \cdots \\
&+ \sum_{n_1=1}^{K}\sum_{\substack{n_2=-K\\ n_2\neq 0}}^{K}\cdots\sum_{\substack{n_d=-K\\ n_d\neq 0}}^{K}\left[\alpha_{n_1\ldots n_d}^{(12\cdots d)}\sin(n_1 z_1 + \cdots + n_d z_d)\right. \\
&\left. + \beta_{n_1\ldots n_d}^{(12\cdots d)}\cos(n_1 z_1 + \cdots + n_d z_d)\right],
\end{aligned}
\tag{32}
$$

where K is the chosen upper bound for all n_i's. Since

$$
\sin(-\mathbf{n}_j^T\mathbf{z}) = -\sin(\mathbf{n}_j^T\mathbf{z}), \qquad \cos(-\mathbf{n}_j^T\mathbf{z}) = \cos(\mathbf{n}_j^T\mathbf{z}),
$$

the functions $\sin(-\mathbf{n}_j^T\mathbf{z})$ and $\cos(-\mathbf{n}_j^T\mathbf{z})$ are combined into $\sin(\mathbf{n}_j^T\mathbf{z})$ and $\cos(\mathbf{n}_j^T\mathbf{z})$, respectively, in Eq. 32. This can be achieved by setting the first summation in each term in Eq. 32 to be over 1 to K, not $-K$ to K.

Equation 32 can be compactly represented as

$$
\mathbf{\Phi x} = \mathbf{y}, \tag{33}
$$

where

$$
\mathbf{\Phi} = \begin{bmatrix}
\psi_1(\mathbf{z}(1)) & \psi_2(\mathbf{z}(1)) & \ldots & \psi_t(\mathbf{z}(1)) \\
\psi_1(\mathbf{z}(2)) & \psi_2(\mathbf{z}(2)) & \ldots & \psi_t(\mathbf{z}(2)) \\
\vdots & \vdots & \ddots & \vdots \\
\psi_1(\mathbf{z}(N)) & \psi_2(\mathbf{z}(N)) & \ldots & \psi_t(\mathbf{z}(N))
\end{bmatrix}
\tag{34}
$$

and

$$\mathbf{x} = \left(\alpha_0\ \alpha_1^{(1)}\ \beta_1^{(1)}\ \ldots\ \beta_{KKKK}^{(1\ldots d)}\right)^T,$$
$$\mathbf{\Phi}_{s*} = \left(1\ \sin(z_1(s))\ \cos(z_1(s))\ \ldots\ \sin\left(K\sum_{k=1}^{d} z_k(s)\right)\ \cos\left(K\sum_{k=1}^{d} z_k(s)\right)\right),$$
$$(s = 1, 2, \ldots, N)$$
$$\mathbf{y} = (\theta(1)\ \theta(2)\ \ldots\ \theta(N))^T.$$

Here, $\mathbf{\Phi}_{s*}$ denotes the sth row of $\mathbf{\Phi}$ and n is the total number of basis functions contained in $\mathbf{\Phi}_{s*}$, and consequently, the total number of unknown parameters contained in $\mathbf{x}$, N is the number of measurements.

Equation 32 can be conveniently represented as a HDMR expansion with terms having different numbers of variables [12,23]:

$$y(\mathbf{z}) = f_0 + \sum_{i=1}^{d} f_i(z_i) + \sum_{1\le i<j\le d} f_{ij}(z_i, z_j) + \sum_{1\le i<j<k\le d} f_{ijk}(z_i, z_j, z_k) + \cdots + f_{12\ldots d}(\mathbf{z}) = f_0 + \sum_{\emptyset\neq u\subseteq d} f_u(\mathbf{z}_u). \tag{35}$$

Using the following conditions (which are approximately [21] or exactly [22] true as required by HE-OD decoding) that for $\mathbf{z} \in [0, 2\pi]^d$

$$\mathbb{E}(\sin(\mathbf{n}^T\mathbf{z})) = 0, \quad \forall\mathbf{n}, \tag{36}$$
$$\mathbb{E}(\cos(\mathbf{n}^T\mathbf{z})) = 0, \quad \forall\mathbf{n}, \tag{37}$$
$$\mathbb{E}(\sin(\mathbf{m}^T\mathbf{z})\cos(\mathbf{n}^T\mathbf{z})) = 0, \quad \forall\mathbf{m}, \mathbf{n} \tag{38}$$
$$\mathbb{E}(\sin(\mathbf{m}^T\mathbf{z})\sin(\mathbf{n}^T\mathbf{z})) = 0, \quad \mathbf{m} \neq \pm\mathbf{n}, \tag{39}$$
$$\mathbb{E}(\cos(\mathbf{m}^T\mathbf{z})\cos(\mathbf{n}^T\mathbf{z})) = 0, \quad \mathbf{m} \neq \pm\mathbf{n}, \tag{40}$$

it is easy to prove that $f_u(\mathbf{z}_u)$ given in Eq. 32 satisfy the characteristic properties of HDMR component functions:

$$\mathbb{E}(f_u(\mathbf{z}_u)) = 0, \quad u \neq \emptyset, \tag{41}$$
$$\mathbb{E}(f_u(\mathbf{z}_u) f_v(\mathbf{z}_v)) = 0, \quad u \neq v, \tag{42}$$

where $\mathbb{E}$ denotes the expected value. Equations 41 and 42 show that the non-constant HDMR component functions have zero mean and all HDMR component functions are mutually orthogonal. Since all non-constant HDMR component functions have zero mean, we have $\mathbb{E}(y(\mathbf{z})) = f_0$ [12,23].

Upon utilizing these properties, a global sensitivity analysis can be performed based on the HDMR expansion [23,24]. Note that

$$\begin{aligned}\mathrm{Var}(y(\mathbf{z})) &= \mathbb{E}[(y(\mathbf{z}) - f_0)^2] = \mathbb{E}[(\sum_{\emptyset \neq u \subseteq d} f_u(\mathbf{z}_u))^2] \\ &= \sum_{\emptyset \neq u \subseteq d} \mathbb{E}[f_u^2(\mathbf{z}_u)] = \sum_{\emptyset \neq u \subseteq d} \mathbb{E}[(f_u(\mathbf{z}_u) - 0)^2] \\ &= \sum_{\emptyset \neq u \subseteq d} \mathrm{Var}(f_u(\mathbf{z}_u)). \end{aligned} \tag{43}$$

The sensitivity indexes can be defined as

$$1 = \frac{\mathrm{Var}(y(\mathbf{z}))}{\mathrm{Var}(y(\mathbf{z}))} = \sum_{\emptyset \neq u \subseteq d} \frac{\mathrm{Var}(f_u(\mathbf{z}_u))}{\mathrm{Var}(y(\mathbf{z}))} = \sum_{\emptyset \neq u \subseteq d} S_u. \tag{44}$$

Since all $S_u \geq 0$ and $\sum_u S_u = 1$, comparing the magnitudes of all S_u's gives the importance order of $\mathbf{z}_u$, which can be also used to identify the mechanism.

3.1.2 Case I: More data than unknown parameters

A quantum-control simulation with four ($d = 4$) selected spectral components of the laser field was performed. A set of 10000 points $\mathbf{z} = (z_1, z_2, z_3, z_4)$ were randomly generated in the range $[0, 2\pi]^4$ and the corresponding signals $y(\mathbf{z}) = \theta(\mathbf{z})$ were calculated. The points $\mathbf{z} = (z_1, z_2, z_3, z_4)$ correspond to spectral components of the field close to the transition frequencies. We set $K = 2$, and altogether there are $n = 625$ unknown parameters $\mathbf{x}$ in Eq. 33 corresponding to possible mechanistic processes. Our goal is to find the key processes in the mechanism, i.e., the sparse solution for $\mathbf{x}$.

One thousand ($\mathbf{z}$, y) points were used as the training data, the remained 9000 data points were used for testing. Since the number (1000) of data is larger than the number (625) of unknown parameter, to construct an underdetermined linear algebraic equation system, first the normal equation for least-squares regression of Eq. 33

$$\mathbf{\Phi}^T \mathbf{\Phi} \mathbf{x} = \mathbf{\Phi}^T \mathbf{y} \tag{45}$$

was obtained where $\mathbf{\Phi}^T \mathbf{\Phi}$ may be nonsingular. Hence, the last 256 equations corresponding to the fourth order HDMR component functions in Eq. 32 were removed to give an underdetermined linear algebraic equation system [14]

$$A\mathbf{x} = \mathbf{b} \tag{46}$$

where A is a 369×625 constant matrix and $\mathbf{b}$ is a 369-dimensional constant vector. Using sD-MORPH regression and setting $\delta = 0.2$, the sparse solution of $\mathbf{x}$ was obtained after 10 recursions. The magnitudes of the unknown parameters, w_i's, obtained from sD-MORPH regression are between $7.13 \times 10^{-2} - 1.78 \times 10^{-7}$. Many

Table 1 The average absolute and relative errors of the HDMR model for training (1000 points) and testing (9000 points) data

Method	Training data		Testing data	
	Abs. err.	Rel. err.	Abs. err.	Rel. err.
sD-MORPH (369 parameters)				
1st order	0.0147	0.2602	0.0151	0.2593
2nd order	0.0094	0.1508	0.0093	0.1475
3rd order	0.0091	0.1532	0.0090	0.1489
4th order	**0.0004**	**0.0073**	**0.0007**	**0.0111**
sD-MORPH (110 parameters)				
1st order	0.0147	0.2602	0.0151	0.2593
2nd order	0.0094	0.1507	0.0093	0.1475
3rd order	0.0091	0.1534	0.0090	0.1489
4th order	**0.0006**	**0.0096**	**0.0007**	**0.0110**
RRLS(369 parameters)	**0.0004**	**0.0073**	**0.0007**	**0.0111**
RRLS(110 parameters)	**0.0006**	**0.0096**	**0.0007**	**0.0110**

parameters of the sparse solution of sD-MORPH regression are not exactly zero, but small numbers. Considering that the values of θ were given to the 3rd digit, therefore, the elements, x_i's, with magnitude less than 10^{-4} were considered as zero and removed to give 110 final nonzero parameters.

The accuracy of the HDMR model obtained by sD-MORPH regression for the training and testing data is given in Table 1. For comparison, the results obtained by RRLS minimization is also provided. The truth plots for training and testing data for the HDMR model are given in Fig. 1.

The sensitivity indexes given in Table 2 were also calculated from the resultant HDMR component functions obtained by sD-MORPH regression with the 1000 training data points.

Analyzing all the results given above, some conclusions can be drawn:

- With 1000 training data the matrix A has row-full rank. The matrix $AW^{-1}A^T$ is nonsingular, and its inverse can be accurately determined. In this case, the results obtained by RRLS minimization and sD-MORPH regression are exactly the same.
- The 4th order HDMR expansion with multi-variate discrete sine and cosine basis functions obtained by sD-MORPH regression is accurate for both training and testing data. Considering that the values of θ are given to the 3rd digit, the 4th order HDMR expansion with 110 parameters is almost exact because the average absolute errors are 0.0006 and 0.0007 for training and testing data, respectively. Thus, the mechanism identified with such an accurate model is reliable.
- The discrete cosine transform (DCT) is a popular method used in signal and image processing, which only uses cosine functions [25] because cosine functions are reported to be much more efficient (fewer functions are needed to approximate a typical signal). However, for the treatment of signals in quantum-control, cosine

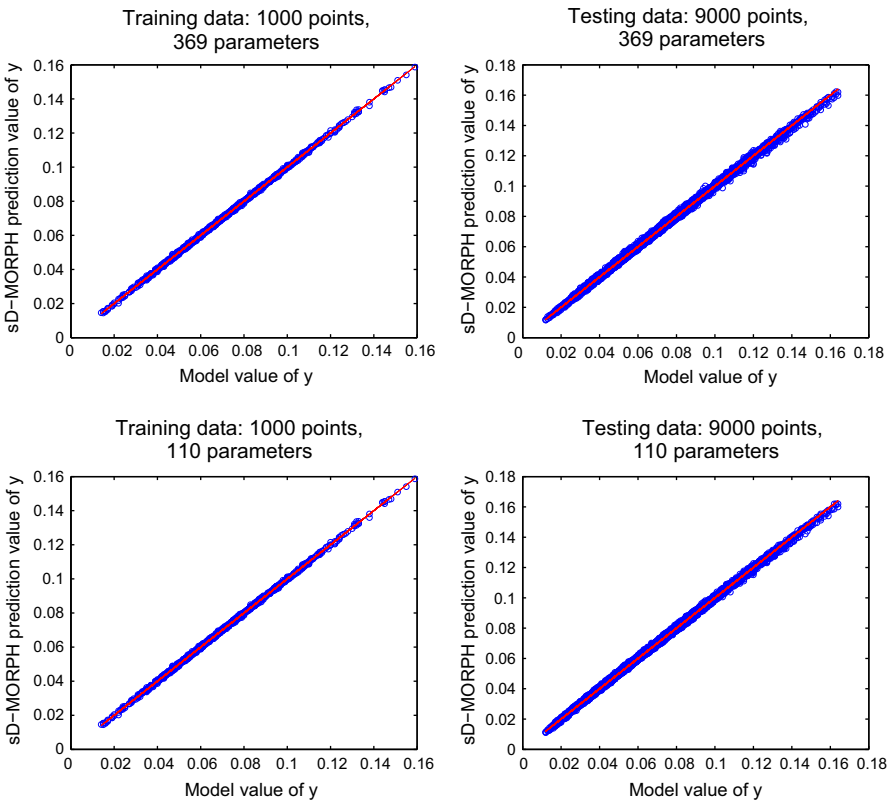

Fig. 1 The truth plots for training (1000 points) and testing (9000 points) data of the HDMR model obtained by sD-MORPH regression for quantum control mechanism analysis

basis functions are not complete. Figure 2 gives the truth plots for training (1000 points) and testing (9000 points) data for the HDMR expansion with only cosine basis functions obtained by sD-MORPH regression.

- The sensitivity indexes provide information about the mechanism. According to the decreasing order of the magnitudes of the sensitivity indexes, the sum of the four most important sensitivity indexes is 0.9338, i.e., only *a few* processes corresponding to the four sensitivity indexes controlling the signal θ.

S_2	S_{23}	S_{1234}	S_{12}	Sum
0.4411	0.2147	0.1512	0.1268	0.9338

- The important processes which control the signal θ may be identified from the magnitude order of the parameters $\{\alpha, \beta\}$ in Eq. 32. Table 3 gives the 10 parameters with the largest magnitudes. For comparison the 10 parameters obtained by HE-OD applied to the data are also listed.
  The largest four parameters (α_0 is always the largest one, and is excluded) with magnitudes larger than 10^{-2} belong to the processes

Table 2 The sensitivity indexes calculated from the HDMR component functions obtained from 1000 training data

Order	Sensitivity index						Sum of each order
1st	S_1	S_2	S_3	S_4			
	0.0291	0.4411	0.0354	0.0002			0.5058
2nd	S_{12}	S_{13}	S_{14}	S_{23}	S_{24}	S_{34}	
	0.1268	0.0022	0.0015	0.2147	0.0007	0.0030	0.3489
3rd	S_{123}	S_{124}	S_{134}	S_{234}			
	0.0067	0.0015	0.0091	0.0039			0.0212
4th	S_{1234}						
	0.1512						0.1512
Total sum							1.0272

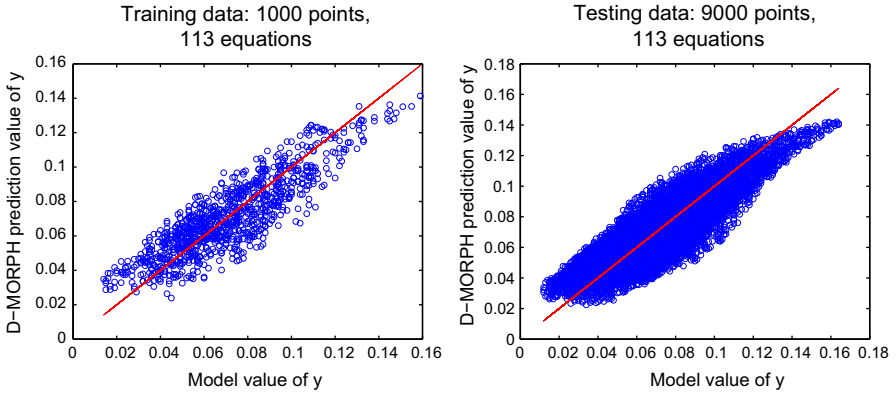

Fig. 2 The truth plots for training and testing data for the HDMR expansion with only cosine basis functions obtained by sD-MORPH regression, thereby indicating the incomplete nature of just cosine basis functions

$$\mathbf{n}_1 = (0\ 1\ 0\ 0),$$
$$\mathbf{n}_2 = (0\ 1\ 1\ 0),$$
$$\mathbf{n}_3 = (1\ -1\ -1\ 1),$$
$$\mathbf{n}_4 = (1\ -1\ 0\ 0).$$

Note that if the coefficient α for $\sin(\mathbf{n}_j^T\mathbf{z})$ is large, then the coefficient β for $\cos(\mathbf{n}_j^T\mathbf{z})$ is also large. Since $\sin(-\mathbf{n}_j^T\mathbf{z})$ and $\cos(-\mathbf{n}_j^T\mathbf{z})$ are included in the terms of $\sin(\mathbf{n}_j^T\mathbf{z})$ and $\cos(\mathbf{n}_j^T\mathbf{z})$, the processes $-\mathbf{n}_j\,(j = 1, 2, 3, 4)$ are also most important. This information identifies the most important processes contributing to the signal θ, which is consistent with the sensitivity indexes S_2, S_{23}, S_{1234}, S_{12}.

Table 3 The importance order of the processes arranged by the decreasing magnitude of the corresponding parameters in Eq. 32 obtained from 1000 training data

Importance order	Basis	Process				Parameters $\{\alpha, \beta\}$	
		n_1	n_2	n_3	n_4	sD-MORPH	HE-OD
1	1	0	0	0	0	7.13×10^{-2}	7.12×10^{-2}
2	cosine	0	1	0	0	2.37×10^{-2}	2.45×10^{-2}
3	cosine	0	1	1	0	1.60×10^{-2}	1.50×10^{-2}
4	cosine	1	−1	−1	1	1.22×10^{-2}	1.12×10^{-2}
5	cosine	1	−1	0	0	1.09×10^{-2}	1.05×10^{-2}
6	sine	1	−1	−1	1	-6.97×10^{-3}	-7.25×10^{-3}
7	cosine	0	0	1	0	6.78×10^{-3}	5.62×10^{-3}
8	sine	1	−1	0	0	5.82×10^{-3}	7.12×10^{-3}
9	sine	0	1	0	0	4.71×10^{-3}	5.42×10^{-3}
10	sine	1	0	0	0	-4.70×10^{-3}	-4.62×10^{-3}

3.1.3 Case II: Less data than unknown parameters

Three hundred $(\mathbf{z}, y)$ points were used as the training data, the remaining 9700 data points were used for testing. Since the number (300) of data is less than the number (625) of unknown parameters, Eq. 33

$$\mathbf{\Phi x} = \mathbf{y}$$

is an underdetermined linear algebraic equation system, where $\mathbf{\Phi}$ is a 300×625 constant matrix. If $\mathbf{y} \in \text{Ran}(\mathbf{\Phi})$ (this is always true when $\mathbf{\Phi}$ has row-full rank), and we may directly use sD-MORPH regression. In this case, RRLS minimization can also apply.

However, no matter whether $\mathbf{y} \in \text{Ran}(\mathbf{\Phi})$ or not, we can always use the normal equation, Eq. 45

$$\mathbf{\Phi}^T \mathbf{\Phi x} = \mathbf{\Phi}^T \mathbf{y}$$

because $\mathbf{\Phi}^T \mathbf{y} \in \text{Ran}(\mathbf{\Phi}^{\text{T}} \mathbf{\Phi})$. Note that the 625×625 constant matrix $\mathbf{\Phi}^T \mathbf{\Phi}$ is singular with rank not larger than 300 (i.e., it does not have row-full rank)

$$\text{rank}(\mathbf{\Phi}^T \mathbf{\Phi}) = \text{rank}(\mathbf{\Phi}) \leq 300,$$

permitting an infinite number of solutions. Moreover, we can use Eq. 46

$$A\mathbf{x} = \mathbf{b}$$

obtained by removing the last 256 equations of the normal equation, where A is an 369×625 constant matrix and has a rank not larger than that of $\mathbf{\Phi}^T \mathbf{\Phi}$. With 300

Table 4 The average absolute and relative errors of the HDMR model for training (300 points) and testing (9700 points) data

Method	Training data		Testing data	
	Abs. err.	Rel. err.	Abs. err.	Rel. err.
sD-MORPH (Eqs. 33, 46) (369 parameters)				
1st order	0.0143	0.2522	0.0151	0.2597
2nd order	0.0095	0.1546	0.0094	0.1483
3rd order	0.0092	0.1564	0.0091	0.1498
4th order	**0.0000**	**0.0000**	**0.0016**	**0.0240**
sD-MORPH (Eqs. 33, 46) (131 parameters)				
1st order	0.0143	0.2523	0.0151	0.2597
2nd order	0.0095	0.1547	0.0094	0.1483
3rd order	0.0093	0.1562	0.0091	0.1498
4th order	**0.0007**	**0.0110**	**0.0015**	**0.0241**
RRLS (Eq. 33)	**0.0000**	**0.0000**	**0.0016**	**0.0240**
RRLS (Eq. 46)	**0.4054**	**6.8982**	**0.4135**	**6.8089**

training data, A does not have row-full rank, but contains all the solutions of Eq. 45 including the sparsest solution [14]. Therefore, when the number of data is less than the number of unknown parameters, both Eqs. 45 and 46 can be used for sD-MORPH regression, but cannot be used for RRLS minimization. In practice, we may need to treat systems not having row-full rank, and sD-MORPH regression is more flexible to search for sparse solutions without any pre-treatment of the system.

Both Eqs. 33 and 46 were used for sD-MORPH regression and RRLS minimization. The accuracy of sD-MORPH regression is exactly the same for Eqs. 33 and 46, only a little reduced compared to that obtained from 1000 training data. The components of x_i with magnitude less than 10^{-4} were considered as zero and removed to give 131 final nonzero components.

Even though the inverse $AW^{-1}A^T$ could be computed, RRLS minimization has a large error for Eq. 46 due to the singular nature of of $AW^{-1}A^T$. In contrast, sD-MORPH regression has the same results for Eqs. 33 and 46, which implies that the singularity of A has no influence upon sD-MORPH regression, and the construction of an underdetermined linear algebraic equation system by removing some equations from the normal equation is general and can be utilized for both the data, either more or less than the unknowns.

The accuracy of the HDMR model for the training and testing data obtained by sD-MORPH regression and RRLS minimization is given in Table 4. The corresponding truth plots are given in Figs. 3 and 4.

The sensitivity indexes and the most important processes determined from the HDMR model obtained by sD-MORPH regression with 300 training data are given in Tables 5 and 6, which are almost the same as those obtained from 1000 training data shown in Tables 2 and 3. This implies that to identify the mechanism much less exper-

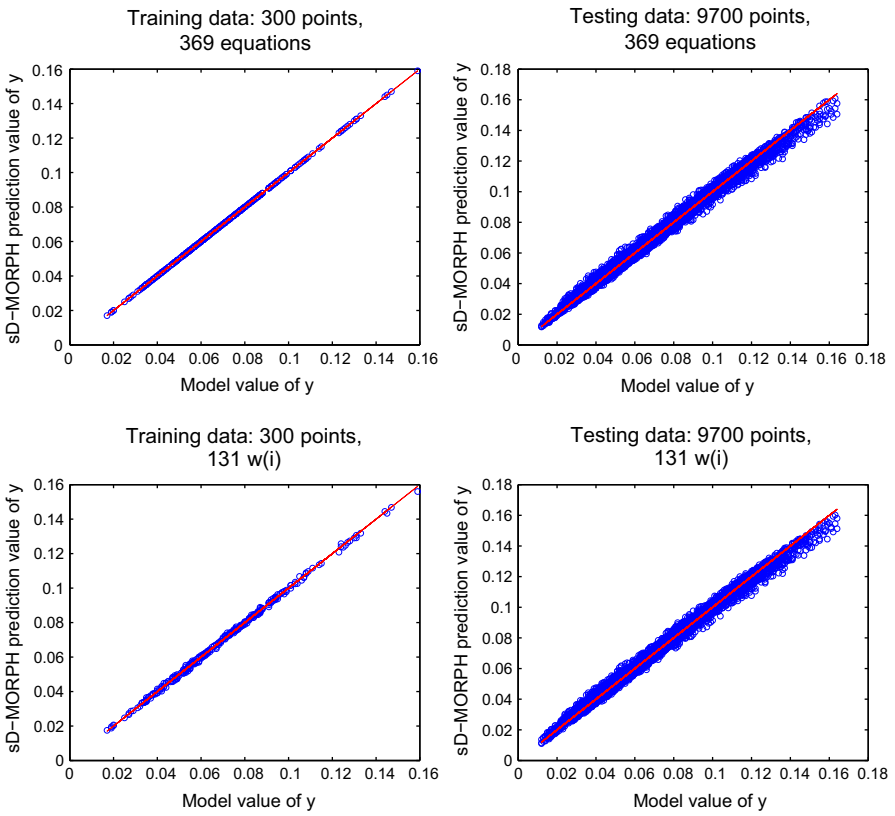

Fig. 3 The truth plots for training (300 points) and testing (9700 points) data of the HDMR model obtained by sD-MORPH regression with Eq. 33 (direct application of the underdetermined linear algebraic equation system) and Eq. 46 (the underdetermined linear algebraic equation system is obtained by removing some equations from the normal equation). The same quality results were obtained

imental data may be needed. This advantage is beneficial for practical experimental applications.

The extracted parameters $\{\alpha, \beta\}$ by sD-MORPH regression are in good agreement with those obtained by the HE-OD technique (see Tables 3, 6), i.e., the processes identified by sD-MORPH regression are consistent with a mechanism dominated by two second order pathways which has been also observed in experiments under similar conditions [22] and confirmed by a separate HE-OD analysis that used the same data set as the one employed in this work.

3.2 Application to composition analysis from a mixture's mass spectrum

In this section, we apply the nnsD-MORPH regression algorithm to a practical problem: composition analysis of an unknown mixture based on using time of flight (TOF) femtosecond (fs) laser-mass spectrometry (MS). We compare the composition MS of

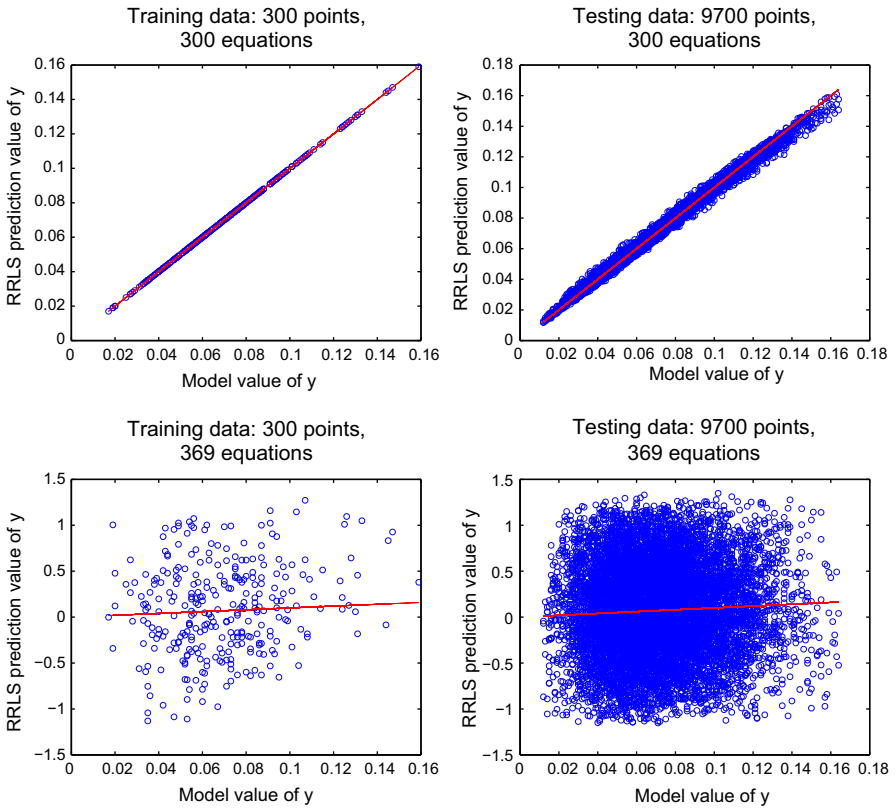

Fig. 4 The truth plots for training (300 points) and testing (9700 points) data of the HDMR model obtained by RRLS minimization with Eq. 33 (where $\mathbf{\Phi}$ has row-full rank) and Eq. 46 (where A does not have row-full rank), respectively. The results demonstrates that RRLS minimization can be used only for the coefficient matrix having row-full rank

Table 5 The sensitivity indexes calculated from the HDMR component functions obtained from 300 training data

Order	Sensitivity index						Sum of each order
1st	S_1	S_2	S_3	S_4			
	0.0298	0.4288	0.0344	0.0004			0.4934
2nd	S_{12}	S_{13}	S_{14}	S_{23}	S_{24}	S_{34}	
	0.1249	0.0023	0.0017	0.2048	0.0006	0.0032	0.3375
3rd	S_{123}	S_{124}	S_{134}	S_{234}			
	0.0072	0.0018	0.0097	0.0035			0.0222
4th	S_{1234}						
	0.1471						0.1471
Total sum							1.0002

Table 6 The importance order of the processes arranged by the decreasing magnitude of the corresponding parameters in Eq. 32 obtained from 300 training data

Importance order	Basis	Process				Parameters $\{\alpha, \beta\}$	
		n_1	n_2	n_3	n_4	sD-MORPH	HE-OD
1	1	0	0	0	0	7.13×10^{-2}	7.12×10^{-2}
2	cosine	0	1	0	0	2.37×10^{-2}	2.45×10^{-2}
3	cosine	0	1	1	0	1.60×10^{-2}	1.50×10^{-2}
4	cosine	1	−1	−1	1	1.23×10^{-2}	1.12×10^{-2}
5	cosine	1	−1	0	0	1.09×10^{-2}	1.05×10^{-2}
6	sine	1	−1	−1	1	-6.91×10^{-3}	-7.25×10^{-3}
7	cosine	0	0	1	0	6.75×10^{-3}	5.62×10^{-3}
8	sine	1	−1	0	0	5.84×10^{-3}	7.12×10^{-3}
9	sine	1	0	0	0	-4.72×10^{-3}	-4.62×10^{-3}
10	sine	0	1	0	0	4.63×10^{-3}	5.42×10^{-3}

a mixture against a well established or user defined spectral database to identify the components in the unknown mixture as well as to quantitatively determine the fractional concentration for each component that is present in the mixture. This section treats the tools introduced in Sects. 1 and 2 with experimental data.

In this test, a total of 11 similar small organic compounds and an extra species (air) are picked to establish a MS data base, which was recorded with the same laser pulse. The MS of a mixture consisting of two species ($CH_2BrCl + CH_2BrI$) with ratio $\sim$1:3 was measured, which was taken as an unknown for test purposes. Since the library size is usually much larger than the number of species in the unknown, the solution can be considered as sparse. Moreover, the quantity of each component is nonnegative. Therefore, nonnegative and sparse nnsD-MORPH regression is used to identify the species present. Figure 5 shows the resultant MS for the 12 compounds and the unknown mixture.

The experimental apparatus includes a fs laser system, a vacuum chamber, and a TOF mass spectrometer. The fs laser system (KMlab, dragon) is capable of generating 790 nm, 1 mJ, sub 30 fs pulses, which are guided and tightly focused inside the vacuum chamber to interact with the gas phase samples introduced through a leak valve. The fragment ions are generated and extracted towards the TOF mass spectrometer detector. Each spectrum is mass calibrated to yield an array of data containing $N = 600$ equally spaced data points as shown in Fig. 5. The separation between two adjacent data points corresponds to half m/q (mass to charge ratio) unit. Note that experimental measurements, calibration and laser alignment cannot be performed exactly the same for all species and unknowns. Therefore, the MS of a species in a pure sample or mixture may have some random differences. But, the MS data were taken under the same experimental conditions, as best as possible, to reduce the data errors.

In order to apply nnsD-MORPH regression, an underdetermined linear algebraic equation system needs to be constructed. It is generally always possible to know which

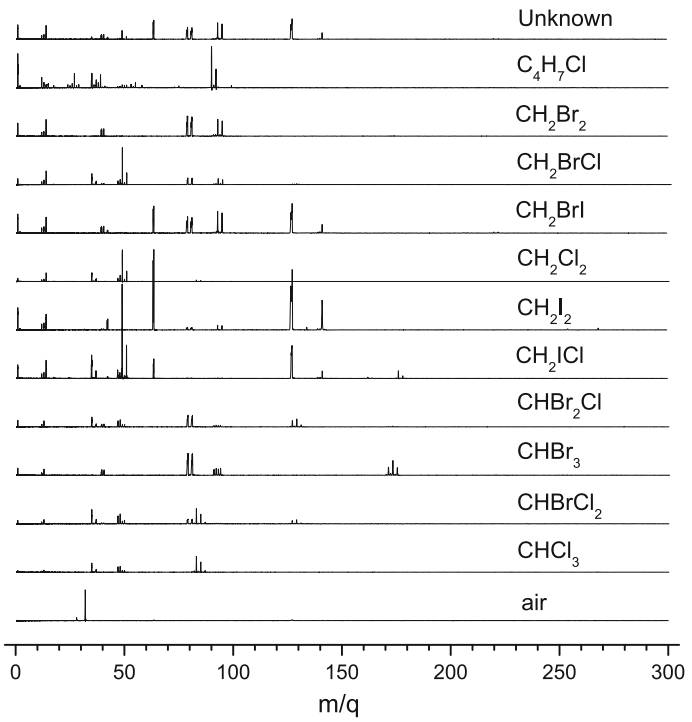

Fig. 5 Mass spectra of 12 reference species and an unknown mixture with the same laser pulse

species in the library is not contained in the unknown mixture. This can be achieved by observing that the mixture's spectrum does not contain characteristic peaks belonging to certain species. For instance, in the above example we are sure that the unknown mixture does not contain air. Then, air is specifically included in the library as a means to construct the underdetermined linear algebraic equation system.

3.2.1 Construction of underdetermined linear algebraic equation system

We first treated the mixture analysis problem by its straight forward nnsD-MORPH procedure. The results will show that a modification of the procedure is needed, as presented in Sect. 3.2.2, in order to obtain reliable performance. Under the assumption that the MS signal $z(k)$ for an unknown mixture at the kth value of m/q, is the superposition of signals of its components at the same kth value of m/q, we have

$$z(k) = \sum_{i=1}^{n} x_i z_i(k) + \varepsilon(k), \quad k = 1, 2, \ldots, N \tag{47}$$

where $z_i(k)$ is the signal of the pure ith ($i = 1, 2, \ldots, n$) component in the library, the coefficient x_i denotes its fraction in the unknown mixture, and $\varepsilon(k)$ is the random error caused by measurement, calibration, laser alignment etc. Suppose that the signals at N distinct values of m/q are measured, then Eq. 47 can be written as

$$\begin{bmatrix} z(1) \\ z(2) \\ \vdots \\ z(N) \end{bmatrix} = \begin{bmatrix} z_1(1) & z_2(1) & \dots & z_n(1) \\ z_1(2) & z_2(2) & \dots & z_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ z_1(N) & z_2(N) & \dots & z_n(N) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} \varepsilon(1) \\ \varepsilon(2) \\ \vdots \\ \varepsilon(N) \end{bmatrix} \tag{48}$$

or in a compact form

$$\mathbf{z} = \left[Z_0 \mid \mathbf{z}_n \right] \mathbf{x} + \epsilon, \tag{49}$$

where $\mathbf{z}$, ϵ and $\mathbf{x}$ are N- and n-dimensional vectors, respectively; Z_0 is an $N \times (n-1)$ constant matrix composed of the first $n-1$ columns of the matrix in Eq. 48 corresponding to the signals of the species in the library which may be contained in the mixture, and $\mathbf{z}_n$ is the last column of the matrix corresponding to the signals of a species for sure not contained in the mixture according to our prior knowledge of the mixture and observation of its MS. For the reference species library given in Fig. 5, $N = 600$, $n = 12$ and the order of $z_i\,(i = 1, 2, \dots, n)$ is chosen to be the order given in the figure such that $\mathbf{z}_n$ is the "signal" of air.

Since $N > n$, to determine the composition vector $\mathbf{x}$ by nnsD-MORPH regression, the normal equation of Eq. 49 for the least-squares regression

$$\begin{bmatrix} Z_0^T Z_0 & Z_0^T \mathbf{z}_n \\ \mathbf{z}_n^T Z_0 & \mathbf{z}_n^T \mathbf{z}_n \end{bmatrix} \mathbf{x} = \begin{bmatrix} Z_0^T \mathbf{z} \\ \mathbf{z}_n^T \mathbf{z} \end{bmatrix} \tag{50}$$

is obtained first. To construct an underdetermined linear algebraic equation system used in nnsD-MORPH regression, the last equation is removed from Eq. 50 to give

$$\left[Z_0^T Z_0 \;\; Z_0^T \mathbf{z}_n \right] \mathbf{x} = Z_0^T \mathbf{z}. \tag{51}$$

The removal of the last equation will not influence the correct determination of mixture composition because we know that the mixture does not contain the last species (air, in the example), and the value of x_n should be zero. Equation 51 can be written as

$$A\mathbf{x} = \mathbf{b}, \tag{52}$$

where A is an $((n-1) \times n)$-rectangular matrix, $\mathbf{b}$ is an $(n-1)$-dimensional vector.

The correct mixture composition can be found by nnsD-MORPH regression if and only if the combination vector $\mathbf{x}$ with some zero elements belongs to the infinite number of solutions in $\mathcal{M}$ for Eq. 52. If not, nnsD-MORPH regression is unable to find the correct solution. As a test, a mixture was constructed with the combination

$$0.25 \quad CH_2BrCl + 0.75 \quad CH_2BrI$$

and the mass spectrum of the artificial mixture is calculated from the the measured MS of CH_2BrCl and CH_2BrI. In this case,

$$\mathbf{x} = (\,0\;0\;0.25\;0.75\;0\;0\;0\;0\;0\;0\;0\;0\,)$$

is a solution of Eq. 52, and nnsD-MORPH regression should give this correct answer. The recursion parameter δ and ε are set to be 0.2 and 10^{-7}, respectively, and 10 recursions were performed. The nnsD-MORPH regression did find the correct sparse solution as above.

Since there exist errors in MS data, *the mass spectrum of a unknown mixture is not an exact superposition of the mass spectra of its components*. The expected solution $\mathbf{x}$ of the mixture (CH_2BrCl:$CH_2BrI = x_3 : x_4 \approx 1 : 3$)

$$\mathbf{x} = (0\ 0 \sim 0.25 \sim 0.75\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)$$

is not a solution in $\mathcal{M}$. Using nnsD-MORPH regression cannot accurately find such a nonnegative sparse solution. Setting $\delta = 0.2$ and $\varepsilon = 10^{-7}$, the solution obtained by nnsD-MORPH regression is

x_1	x_2	x_3	x_4	x_5	x_6
0.001	0.005	0.182	0.704	0.019	0.001
x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
0.009	−0.073	0.006	0.001	0.033	−0.004

The problem of treating error contaminated data identified here is addressed in the section below.

3.2.2 Sequential determination of nonnegative and sparse solution for mixture composition

To correctly determine the composition of an unknown mixture, a sequential determination procedure is employed. Instead of determining the single most sparse solution, we find the best solutions for each given level of sparsity. The basic concept behind of the modification for nnsD-MORPH regression is based on the following observations:

- When mass spectral data have errors, nnsD-MORPH regression cannot readily identity the correct sparse solution. Since A in Eq. 52 is an $(m \times n)$-rectangular matrix (in the current case, $m = n - 1$), the solution $\mathbf{x}$ can have a sparsity m with $n - m$ (in the current case, 1) small numbers which might be zero when the mass spectra data *do not* have errors.
- The small numbers in $\mathbf{x}$ reflecting various data errors suggest that the corresponding species are likely to be absent in the mixture. If so, removing these species from the reference species library will *not significantly change* the combination coefficients of the other remaining species obtained from a reduced species library. Thus, these species deemed irrelevant are removed step-by-step. In each step, the species with the smallest magnitude of combination coefficients is removed. If removing a species causes significant changes of the remaining combination coefficients determined from the reduced species library, this species is not considered

irrelevant at least at this point in the procedure and will be retained. In the overall process, the species air (known to not be present) is always kept to construct the underdetermined linear algebraic equation system, Eq. 52.

- When there is no remaining species to remove, i.e., the magnitudes of all nonzero combination coefficients are not smaller than a chosen threshold (depending on an estimate of the experimental errors), the removal process stops and the final sparsest solution is obtained.

The sequential data analysis procedure for the test case is given below. The first species to be removed from the library is CH_2I_2 with the smallest magnitude $x_6 = 0.0006$. Then, the library contains 11 species. After removing the last equation for air, we obtain a (10×11)-dimensional matrix A (i.e., $m = 10$). $CHBrCl_2$ with the smallest magnitude $x_{10} = 0.0001$ in the solution is the next removed species. The process of removing a species from the library continues until $m = 4$. The removed species are in the order of $x_6, x_{10}, x_1, x_2, x_9, x_7, x_{11}$. Table 7 gives the resultant combination coefficients for $m = 11, 10, \ldots, 4$. The average absolute and relative errors of the signals for the mixture predicted by the nnsD-MORPH regression solutions with $m = 11, 10, \ldots, 4$ are given in Table 8. All combination coefficients should be nonnegative, but some combination coefficients with small magnitudes are still negative which implies that the largest shrinkage weight $1/\varepsilon$ setting for negative x_i (c_i has the maximum value $1/\varepsilon$ when setting $x_i^{(l)} = 0$ for negative $x_i^{(l)}$) may not be large enough.

The truth plots for the signals of the mixture's mass spectrum predicted by nnsD-MORPH regression with $m = 11$ and $m = 4$ are given in Fig. 6. The accuracy for m from 11 to 4 is almost the same. Since the remaining combination coefficients and the average absolute and relative errors in the whole process do not change significantly, we may draw the conclusion: *all removed species are not contained in the mixture*.

Considering the data error (the signal $z(k)$ is given to the 3 digit), the difference between the magnitudes of $x_5(0.0576)$ and $x_8(-0.0563)$ in Table 7 with $m = 4$ is not significant, and either one may be removed first in the next iterative step. When x_5 is removed first, the resultant combination coefficients determined in the following reduced species libraries and the corresponding average absolute and relative errors for $m = 4, 3, 2$ are given in Tables 9 and 10.

From Table 9 we see that after removing CH_2Cl_2 (x_5) from $m = 4$, the remaining combination coefficients in $m = 3$ do not contain qualitative changes. Therefore, $CH_2Cl_2(x_5)$ may be considered as an irrelevant species. Similarly, $CHBr_2Cl$ (x_8) was removed from $m = 3$ to give a solution with two species CH_2BrCl ($x_3 = 0.2052$) and CH_2BrI ($x_4 = 0.6987$). Since 0.2052 and 0.6987 are not small, it is not reasonable to further remove any species, and the final sparsest solution has been obtained. Figure 7 gives the truth plot for the mixture's signal given by the solution of nnsD-MORPH regression with CH_2BrCl and CH_2BrI only. The ratio $x_4/x_3 = 3.4050$ is close to the experiment value.

As another test of the procedure, instead of the last two sequential steps above, we start by removing $CHBr_2Cl$ (x_8) first from $m = 4$. The results are given in Tables 11 and 12.

Table 7 The solutions of sequential determination by nnsD-MORPH regression for $m = 11, 10, \ldots, 4$ library members (without air)

Compound		$m = 11$	$m = 10$	$m = 9$	$m = 8$	$m = 7$	$m = 6$	$m = 5$	$m = 4$
C_4H_7Cl	x_1	−0.0007	−0.0007	−0.0007	**0.0000**	0.0000	0.0000	0.0000	0.0000
CH_2Br_2	x_2	0.0046	0.0033	0.0033	0.0031	**0.0000**	0.0000	0.0000	0.0000
CH_2BrCl	x_3	0.1819	0.1808	0.1809	0.1813	0.1858	0.1927	0.1775	0.1721
CH_2BrI	x_4	0.7041	0.7059	0.7058	0.7058	0.7073	0.7071	0.7111	0.7112
CH_2Cl_2	x_5	0.0185	0.0207	0.0207	0.0203	0.0194	0.0130	0.0500	0.0576
CH_2I_2	x_6	0.0006	**0.0000**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CH_2ICl	x_7	0.0090	0.0087	0.0087	0.0087	0.0072	0.0066	**0.0000**	0.0000
$CHBr_2Cl$	x_8	−0.0727	−0.0719	−0.0720	−0.0719	−0.0722	−0.0641	−0.0603	−0.0563
$CHBr_3$	x_9	0.0064	0.0063	0.0063	0.0062	0.0064	**0.0000**	0.0000	0.0000
$CHBrCl_2$	x_{10}	0.0009	0.0001	**0.0000**	0.0000	0.0000	0.0000	0.0000	0.0000
$CHCl_3$	x_{11}	0.0325	0.0325	0.0327	0.0317	0.0311	0.0227	0.0128	**0.0000**
Air	x_{12}	−0.0041	−0.0049	−0.0025	−0.0024	−0.0024	−0.0010	−0.0013	−0.0008

The $x_i^{(l)}$ value for the removed species is set to be zero in the bold face

Table 8 The average absolute and relative errors of the signals for the mixture's MS predicted by nnsD-MORPH regression with $m = 11, 10, \ldots, 4$ library members (without air)

m	11	10	9	8	7	6	5	4
Abs.	0.009	0.009	0.009	0.009	0.009	0.009	0.010	0.010
Rel.	0.137	0.173	0.174	0.174	0.173	0.171	0.178	0.174

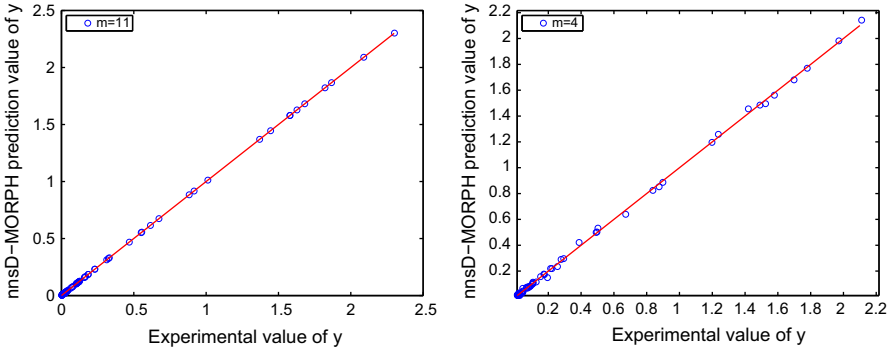

Fig. 6 Comparison of truth plots for the mixture's MS signals predicted by nnsD-MORPH regression with $m = 11$ and $m = 4$. No significant difference can be found

Table 9 The solutions obtained by nnsD-MORPH regression for $m = 4, 3, 2$ starting with removing CH_2Cl_2 (x_5)

Compound		$m = 4$	$m = 3$	$m = 2$
C_4H_7Cl	x_1	0.0000	0.0000	0.0000
CH_2Br_2	x_2	0.0000	0.0000	0.0000
CH_2BrCl	x_3	0.1721	0.2203	0.2052
CH_2BrI	x_4	0.7112	0.7073	0.6987
CH_2Cl_2	x_5	0.0576	**0.0000**	0.0000
CH_2I_2	x_6	0.0000	0.0000	0.0000
CH_2ICl	x_7	0.0000	0.0000	0.0000
$CHBr_2Cl$	x_8	−0.0563	−0.0661	**0.0000**
$CHBr_3$	x_9	0.0000	0.0000	0.0000
$CHBrCl_2$	x_{10}	0.0000	0.0000	0.0000
$CHCl_3$	x_{11}	**0.0000**	0.0000	0.0000
Air	x_{12}	−0.0008	−0.0002	0.0000

Table 10 The average absolute and relative errors of nnsD-MORPH regression solutions with $m = 4, 3, 2$ starting with removing CH_2Cl_2 (x_5)

	Remaining species		
	x_3, x_4, x_5, x_8	x_3, x_4, x_8	x_3, x_4
Abs. error	0.010	0.010	0.016
Rel. error	0.174	0.175	0.244

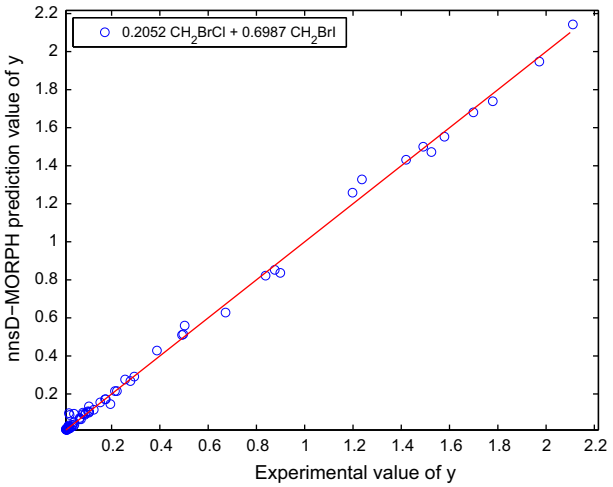

Fig. 7 The truth plot for the nonnegative sparse solution with CH_2BrCl and CH_2BrI only

Table 11 The solutions obtained by nnsD-MORPH regression for $m = 4, 3, 2$ starting from removing $CHBr_2Cl$ (x_8)

Compound		$m = 4$	$m = 3$	$m = 2$
C_4H_7Cl	x_1	0.0000	0.0000	0.0000
CH_2Br_2	x_2	0.0000	0.0000	0.0000
CH_2BrCl	x_3	0.1721	0.1010	**0.0000**
CH_2BrI	x_4	0.7112	0.7103	0.7219
CH_2Cl_2	x_5	0.0576	0.1307	0.2486
CH_2I_2	x_6	0.0000	0.0000	0.0000
CH_2ICl	x_7	0.0000	0.0000	0.0000
$CHBr_2Cl$	x_8	−0.0563	**0.0000**	0.0000
$CHBr_3$	x_9	0.0000	0.0000	0.0000
$CHBrCl_2$	x_{10}	0.0000	0.0000	0.0000
$CHCl_3$	x_{11}	**0.0000**	0.0000	0.0000
Air	x_{12}	−0.0008	0.0002	0.0000

Table 12 The average absolute and relative errors of nnsD-MORPH regression solutions with $m = 4, 3, 2$ starting from removing $CHBr_2Cl$ (x_8)

	Remaining species		
	x_3, x_4, x_5, x_8	x_3, x_4, x_5	x_4, x_5
Abs. error	0.010	0.013	0.015
Rel. error	0.174	0.206	0.197

Table 11 shows that removing x_8 caused some qualitative change: x_5 becomes the second important component increasing significantly from 0.0576 to 0.1307. This implies that we cannot consider $CHBr_2Cl$ (x_8) as an irrelevant species in this application of the scheme. Therefore, its removal is not proper even though the resultant solution with 3 species (CH_2BrCl (0.1010), CH_2BrI (0.7103) and CH_2Cl_2 (0.1307))

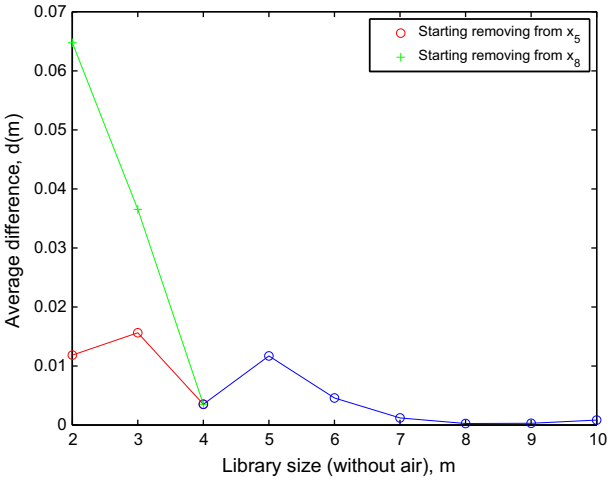

Fig. 8 Changes of average difference $d(m)$ between the nonzero elements given by nnsD-MORPH regression before and after removing one species from the library

has good accuracy for the mixture's signal, and we still cannot consider it as the correct final solution. If we further remove CH_2BrCl (x_3) (this is improper because $x_3 = 0.1010$ is not a small number and we cannot consider it as an irrelevant species), the resultant solution with two species CH_2BrI (0.7219) and CH_2Cl_2 (0.2486) has an even better accuracy than that for the mixture CH_2BrCl (0.2052) and CH_2BrI (0.6987), but this solution is not proper.

Figure 8 gives the changes of the quantity $d(m)$ in the removal process defined as

$$d(m) = \sum_{\substack{i=1 \\ x_i(m)\neq 0}}^{12} |x_i(m+1) - x_i(m)|/(m+1), \qquad m = 10, 9, \ldots, 2, \tag{53}$$

where $\mathbf{x}(m)$ is the solution of nnsD-MORPH regression obtained from the library with $m+1$ species. $d(m)$ represents the average difference between the nonzero elements of $\mathbf{x}$ obtained by nnsD-MORPH regression before and after removing one species. The $d(m)$ does not change much before $m = 4$ which implies that removing one species from the library does not significantly change the values of the combination coefficients for the remaining species. After $m = 4$, however, starting from removal of $CHBr_2Cl$ (x_8) there is a significant increase of $d(m)$, which implies that at this step in the process the removal of $CHBr_2Cl$ (x_8) is improper. In contrast, starting with removing CH_2Cl_2 (x_5) does not significantly increase $d(m)$, which leads to the correct two component solution (CH_2BrCl (0.2052) and CH_2BrI (0.6987)).

The algorithm for mixture analysis, Knowitall [26], analyzed the test data set and gave possible solutions arranged according to fitting accuracy shown in Table 13. The accuracy measure for Knowitall is referred to as the Hit Quality Index (HQI), and the larger the better. The accuracy measure for the nnsD-MORPH regression is the averaged absolute error, and the smaller the better.

Table 13 The comparison of the solutions given by algorithm Knowitall and nnsD-MORPH regression

Components	Combination coefficients		Accuracy measure	
	Knowitall	nnsD-MORPH	Knowitall	nnsD-MORPH
(CH_2BrCl, CH_2BrI, CH_2Cl_2)	(0.13, 0.76, 0.12)	(0.10, 0.71, 0.13)	973.37	0.0130
(CH_2BrCl, CH_2BrI)	(0.23, 0.77)	(0.21, 0.70)	969.12	0.0155
(CH_2BrI, CH_2Cl_2)	(0.77, 0.23)	(0.72, 0.25)	968.14	0.0145

Table 13 shows that the combination coefficients and accuracy measure obtained by the two methods are very consistent. However, if accuracy measure is the only consideration, we would arrive the wrong solution $CH_2BrCl+CH_2BrI+CH_2Cl_2$. Sequentially removing irrelevant species from the reference species library can help nnsD-MORPH regression to correctly determine the mixture composition (i.e., CH_2BrCl + CH_2BrI, as explained in the procedure leading to Fig. 8). The criterion for identification of irrelevant species is simple and efficient. However, if the mixture contains some species in very small amounts, especially at the noise level in the data, it may be difficult to correctly identify whether it is an irrelevant species, and the sequential removal algorithm may fail to correctly determine the mixture composition. Nevertheless, sequential determination by nnsD-MORPH can still give the combination solution with the best accuracy for different sparsity.

4 Conclusion

In this paper, D-MORPH regression has been extended to determine *sparse* and *nonnegative sparse* solutions. The weight used in RRLS minimization is adopted and modified in the construction of the cost function $\mathcal{K}$ for D-MORPH regression to determine sparse and nonnegative sparse solutions. Compared to IRLS and RRLS, the advantage of sparse and nonnegative sparse D-MORPH regression is that the matrix $\mathbf{\Phi}$ of an underdetermined linear algebraic equation system does not need to have a row-full rank, which makes D-MORPH regression more flexible in searching for sparse solutions in practical applications. Simulation data for quantum-control-mechanism identification as well as experimental mass spectral data for determining the composition of an unknown mixture of chemical species were used as illustrative examples.

For composition analysis of a mixture's MS, due to experimental errors, the correct combination coefficients does not belong to the solutions in $\mathcal{M}$ and cannot be directly determined by nnsD-MORPH regression. To solve this problem, the best solution for each given sparsity is searched for in a sequential fashion instead of determining a single sparsest solution. This can be achieved by reducing one species from the reference library step-by-step, until the stable correct solution is obtained. The same procedure could be applied to sparse mixture identification with other type of spectral data.

Acknowledgments Support for this work was provided by ONR with Account Number N00014-11-1-0716.

References

1. I. Daubechies, R. Devore, M. Fornasier, C.S. Güntürk, Commun. Pure Appl. Math. **LXIII**, 0001–0038 (2010)
2. A.E. Yagle, Non-iterative reweighted-norm least-squares local ℓ_0 minimization for sparse solutions to underdetermined linear system of equations, http://web.eecs.umich.edu/~aey/sparse/sparse11.pdf
3. K. Mohan, M. Fazel, J. Mach. Learn. Res. **13**, 3441–3473 (2012)
4. M. Fornasier, H. Rauhut, R. Ward, Low-rank matrix recovery via iteratively reweighted least squares minimization, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.174.8001
5. H. Firouzi, M. Farivar, M. Babaie-Zadeh, C. Jutten, Approximate sparse decomposition based on smoothed ℓ_0-Norm, http://arxiv.org/pdf/0811.2868
6. R.G. Baraniuk, Compressive sensing [Lecture Notes]. IEEE Signal Process. Mag. **24**(4), 118–121 (2007)
7. E.J. Candés, Compressive Sampling, in *International Congress of Mathematicians*, Vol. III, pp. 1433–1452 (2006). European Mathematical Society, Zürich
8. D.L. Donoho, Y. Tsaig, IEEE Trans. Inf. Theory **54**(11), 4789–4812 (2008)
9. Y. Li, SIAM J. Optim. **3**(3), 609–629 (1993)
10. H. Rabitz, O.F. Alis, J. Math. Chem. **25**, 197–233 (1999)
11. G. Li, C. Rosenthal, H. Rabitz, J. Phys. Chem. A **105**(33), 7765–7777 (2001)
12. G. Li, H. Rabitz, J. Math. Chem. **50**, 99–130 (2012)
13. G. Li, H. Rabitz, J. Math. Chem. **48**, 1010–1035 (2010)
14. G. Li, R. Rey-de-Castro, H. Rabitz, J. Math. Chem. **50**, 1747–1764 (2012)
15. G. Li, C. Bastian, W. Welsh, H. Rabitz, J. Phys. Chem. A (in press)
16. C.R. Rao, S.K. Mitra, *Generalized Inverse of Matrix and Its Applications* (Willey, New York, 1971)
17. F.I. Kushnirskii, M.E. Primak, Regression with nonnegative coefficients. Kibernetika (Russian) **1**, 151–152 (1973)
18. C. Brif, R. Chakrabarti, H. Rabitz, Adv. Chem. Phys. **148**, 1 (2012)
19. R.S. Judson, H. Rabitz, Phys. Rev. Lett. **68**(10), 1500–1503 (1992)
20. A. Mitra, H. Rabitz, Phys. Rev. A **67**(3), 33407 (2003)
21. R. Rey-de-Castro, H. Rabitz, Phys. Rev. A **81**, 063422 (2010)
22. R. Rey-de-Castro, Z. Leghtas, H. Rabitz, Phys. Rev. Lett. **110**, 223601 (2013)
23. G. Li, H. Rabitz, J. Math. Chem. **52**, 2052–2073 (2014)
24. G. Li et al., J. Phys. Chem. A **114**, 6022–6032 (2010)
25. N. Ahmed, T. Natarajan, K.R. Rao, Discrete cosine transform. IEEE Trans. Comput. **C23**(1), 90–93 (1974)
26. http://www.knowitall.com/