"Supporting Information"

# Explicit-Implicit Mapping Approach to Nonlinear Single-mixture Blind Separation of Sparse Nonnegative Sources: Pure Components Extraction from Nonlinear Mixture Mass Spectra

Ivica Kopriva<sup>1\*</sup>, Ivanka Jerić<sup>2\*</sup> and Lidija Brkljačić<sup>2</sup>

Ruđer Bošković Institute, Bijenička cesta 54, HR-10000, Zagreb, Croatia

<sup>1</sup>Division of Laser and Atomic Research and Development

<sup>2</sup>Division of Organic Chemistry and Biochemistry

<sup>1\*</sup>Phone: +385.1.4571.286. Fax: +385.1.4680.104. E-mail: ikopriva@irb.hr.

<sup>2\*</sup> Phone: +385.1.4560.998. Fax: +385.1.4680.195. E-mail: ijeric@irb.hr.

# Table of contents:

Suppression of higher order (error) terms - reproduced section 2.3 from reference 18.3-6Description of chemical reactions.7-9Mass spectrometry measurements.10Figure S-1. Structures of possible components present in first reaction mixture.11Table S-1. First nonlinear reaction. Normalized cross-correlation coefficients between 25 pure coefficient above 0.1.7Figure S-2. Experimental study. 9 chromatograms recorded during first nonlinear chemical reaction of peptide synthesis.13-17Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.37-49NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.63Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized correlation coefficient above 0.1.63Figure S-3. Structures of possible components present in second reaction mixture.64	Subject	Page			
Description of chemical reactions.7-9Mass spectrometry measurements.10Figure S-1. Structures of possible components present in first reaction mixture.11Table S-1. First nonlinear reaction. Normalized cross-correlation coefficients between 25 pure components generated in. Pairs of pure components are identified with normalized correlation coefficient above 0.1.7Figure S-2. Experimental study. 9 chromatograms recorded during first nonlinear chemical reaction of peptide synthesis.13-17Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.37-49NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized correlation coefficient above 0.1.63Figure S-3. Structures of possible components present in second reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.63Figure S-4. Structures of possible components present in third reaction mixture.64Table S-3. Structures of possible components present in third reaction mixture.64	Suppression of higher order (error) terms - reproduced section 2.3 from reference 18.	3-6			
Mass spectrometry measurements.10Figure S-1. Structures of possible components present in first reaction mixture.11Table S-1. First nonlinear reaction. Normalized cross-correlation coefficients between 25 pure components generated in. Pairs of pure components are identified with normalized correlation coefficient above 0.1.7Figure S-2. Experimental study. 9 chromatograms recorded during first nonlinear chemical reaction of peptide synthesis.13-17Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Game S-7. Structures of possible components present in second reaction mixture.63Figure S-8. Structures of possible components present in third reaction. Pairs of pure components are identified with normalized crorelation coefficient above 0.1.64Figure S-8. Third nonlinear react	Description of chemical reactions.	7-9			
Figure S-1. Structures of possible components present in first reaction mixture.11Table S-1. First nonlinear reaction. Normalized cross-correlation coefficients between 25 pure components generated in. Pairs of pure components are identified with normalized correlation coefficient above 0.1.7Figure S-2. Experimental study. 9 chromatograms recorded during first nonlinear chemical reaction of peptide synthesis.13-17Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- MU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Game S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between a up on ponents generated in non	Mass spectrometry measurements.	10			
Table S-1. First nonlinear reaction. Normalized cross-correlation coefficients between 25 pure components generated in. Pairs of pure components are identified with normalized correlation coefficient above 0.1.7Figure S-2. Experimental study. 9 chromatograms recorded during first nonlinear chemical reaction of peptide synthesis.13-17Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Game S-2. Second nonlinear reaction. Normalized cross-correlation coefficient above 0.1.64Figure S-3. Structures of possible components present in third reaction mixture.64Carbon S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 2.2.65Carbon S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 2.3. Third nonlinear reaction. Normalized cross-correlation co	Figure S-1. Structures of possible components present in first reaction mixture.	11			
components generated in. Pairs of pure components are identified with normalized correlation coefficient above 0.1.13-17Figure S-2. Experimental study. 9 chromatograms recorded during first nonlinear chemical reaction of peptide synthesis.13-17Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- to generated in nonlinear chemical reaction mixture.63Gigure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- to generated assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.63Figure S-7. Structures of possible components present in second reaction mixture.63Gamee S-8. Structures of possible components present in third reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between a nonlinear chemical reaction mixture.64Gamee S-8. Structures of possible components present in third reaction mixture.64<	Table S-1. First nonlinear reaction. Normalized cross-correlation coefficients between 25 pure	7			
coefficient above 0.1.Figure S-2. Experimental study. 9 chromatograms recorded during first nonlinear chemical reaction of peptide synthesis.13-17Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Game S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.64A Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between a nonlinear reac	components generated in. Pairs of pure components are identified with normalized correlation				
Figure S-2. Experimental study. 9 chromatograms recorded during first nonlinear chemical reaction of peptide synthesis.13-17Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.37-49Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Game S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.646528 nure components generated in nonlinear chemical reaction Pairs of pure	coefficient above 0.1.				
reaction of peptide synthesis.Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.37-49Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.63Figure S-7. Structures of possible components present in second reaction mixture.63Game S-8. Structures of possible components present in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.64Able S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between a use components generated in nonlinear chemical reaction pairs of pure components generated in nonlinear chemical reaction pairs of pure	Figure S-2. Experimental study. 9 chromatograms recorded during first nonlinear chemical	13-17			
Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full integration of chromatograms shown in Figure S-1.18-22Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.37-49Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- tat yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- tat yielded best result are also reported.50-62Figure S-7. Structures of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.63Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.64Calle S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between are identified with normalized cross-correlation coefficients between are identified in nonlinear chemical reaction mixture.64	reaction of peptide synthesis.				
integration of chromatograms shown in Figure S-1.23-36Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.37-49Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized correlation coefficient above 0.1.63Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between are identified with normalized cross-correlation coefficients between the pure components generated in nonlinear chemical reaction mixture.65	Figure S-3. Experimental study. Mass spectra of 9 nonlinear mixtures obtained by full	18-22			
Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the first reaction.23-36Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.37-49Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized correlation coefficient above 0.1.63Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between a pure components generated in nonlinear chemical reaction mixture.64	integration of chromatograms shown in Figure S-1.				
first reaction.Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.37-49Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized correlation coefficient above 0.1.63Figure S-8. Structures of possible components present in third reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 28 pure components generated in nonlinear chemical reaction mixture.65	Figure S-4. Experimental study. Mass spectra of 25 pure components that participate in the	23-36			
Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM- NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.37-49Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized correlation coefficient above 0.1.63Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 28 pure components generated in nonlinear chemical reaction Pairs of pure65	first reaction.				
NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized cross-correlation coefficients between 16464Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 28 pure components generated in nonlinear chemical reaction Pairs of pure65	Figure S-5. Experimental study. Mass spectra of 25 components separated by PTs-EKM-	37-49			
component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized cross-correlation coefficients between 0.1.64Figure S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 28 pure components generated in nonlinear chemical reaction pairs of pure65	NMU algorithm and assigned to pure components from the library. In addition to pure				
that yielded best result are also reported.50-62Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 28 pure components generated in nonlinear chemical reaction pairs of pure65	component index, value of the normalized correlation coefficient and preprocessing transform				
Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM- EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.50-62Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.64Components are identified with normalized cross-correlation coefficients between65Output656528 pure components generated in nonlinear chemical reaction64	that yielded best result are also reported.				
EKM-NMU algorithm and assigned to pure components from the library. In addition to pure component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.63Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.64Components are identified with normalized cross-correlation coefficients between a pure components generated in nonlinear chemical reaction mixture.64Calle S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between a pure components generated in nonlinear chemical reaction pairs of pure65	Figure S-6. Experimental study. Mass spectra of 25 components separated by PTs-EFM-	50-62			
component index, value of the normalized correlation coefficient and preprocessing transform that yielded best result are also reported.63Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.63Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 28 pure components generated in nonlinear chemical reaction Pairs of pure	EKM-NMU algorithm and assigned to pure components from the library. In addition to pure				
that yielded best result are also reported.63Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients63between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 28 pure components generated in nonlinear chemical reaction Pairs of pure	component index, value of the normalized correlation coefficient and preprocessing transform				
Figure S-7. Structures of possible components present in second reaction mixture.63Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.63Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 28 pure components generated in nonlinear chemical reaction Pairs of pure	that yielded best result are also reported.				
Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients63between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between28 pure components generated in nonlinear chemical reaction Pairs of pure	Figure S-7. Structures of possible components present in second reaction mixture.	63			
between 19 pure components generated in nonlinear chemical reaction. Pairs of pure components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between 28 pure components generated in nonlinear chemical reaction Pairs of pure65	Table S-2. Second nonlinear reaction. Normalized cross-correlation coefficients	63			
components are identified with normalized correlation coefficient above 0.1.64Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between6528 pure components generated in nonlinear chemical reaction Pairs of pure65	between 19 pure components generated in nonlinear chemical reaction. Pairs of pure				
Figure S-8. Structures of possible components present in third reaction mixture.64Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between6528 pure components generated in nonlinear chemical reaction Pairs of pure65	components are identified with normalized correlation coefficient above 0.1.				
<b>Table S-3.</b> Third nonlinear reaction. Normalized cross-correlation coefficients between6528 pure components generated in nonlinear chemical reaction Pairs of pure	Figure S-8. Structures of possible components present in third reaction mixture.	64			
28 pure components generated in nonlinear chemical reaction Pairs of pure	Table S-3. Third nonlinear reaction. Normalized cross-correlation coefficients between				
- point to an point and an	28 pure components generated in nonlinear chemical reaction. Pairs of pure				
components are identified with normalized correlation coefficient above 0.1.	components are identified with normalized correlation coefficient above 0.1.				

#### Suppression of higher order (error) terms - reproduced section 2.3 from reference 18.

Mass spectra of 25 pure components recorded in nonlinear chemical reaction of peptide bond formation are shown in Figure S-4 in Supporting Information, illustrate diversity of morphologies. Some have few very dominant (large) peaks (see spectra of pure components 1, 2, 8, 13, 16, 17, 18, 19, 20, 21, 22, 23, 24 and 25), some have intensities distributed on several m/z values, whereas intensities can be small (see spectra of pure components 3, 4, 5, 6, 7, 9, 10, 11, 12, 14 and 15). It is thus hard to propose one preprocessing (thresholding) transform for suppression of higher order terms induced by nonlinear mixing process. We, therefore, propose the combination of methods for this purpose.

### Robust principal component analysis

RPCA has been proposed in [S1, S2] to decompose data vector  $\mathbf{x}$  in (4) into sum of two vectors:  $\mathbf{x}=\mathbf{a}+\mathbf{e}$ . This problem is a special case of matrix decomposition problem:  $\mathbf{X}=\mathbf{A}+\mathbf{S}$ . Provided that  $\mathbf{A}$  is low rank matrix and  $\mathbf{E}$  is sparse matrix decomposition is unique and it is obtained as a solution of the optimization problem:

minimize 
$$\|\mathbf{A}\|_{*} + \lambda \|\mathbf{E}\|_{1}$$
 subject to:  $\mathbf{A} + \mathbf{E} = \mathbf{X}$ . (S1)

Thereby,  $\|\mathbf{A}\|_* = \sum_{i=1}^{I \le N} \sigma_i$  denotes nuclear norm (sum of singular values) and  $I \le N$  is a rank of matrix

**A**;  $\|\mathbf{E}\|_{1} = \sum_{n=1}^{N} \sum_{t=1}^{T} e_{nt}$  denotes  $\ell_{1}$ -norm of **E** and  $\lambda \approx 1/\sqrt{T}$  is a regularization constant. In case of a

vector nuclear norm is equivalent to  $\ell_2$ -norm and problem (S2) relates to minimization of  $\ell_2$ norm of **a**. In term of equivalent linear BSS problem (4), **a** is associated with first and second order terms and **e** is associated with *HOT*. **a** is actually represented by linear mixture model composed of 2M + M(M-1)/2 sources and *I* mixture. **e** is comprised of monomials (products of the original source components) of the order three- or higher. Since by assumption A4 source components are sparse in support and amplitude their three- and higher-order products are either zero or very small. Thus, **e** is sparse. Therefore, it is justified to use RPCA decomposition of **x** in (4) to suppress higher-order terms induced by nonlinear mixing process. That yields approximation of **x**, that is **a** in eq.(6), with suppressed higher-order terms. In the experiments reported in Section 3 we have used accelerated proximal gradient algorithm [S3], available with a MATLAB code at [S4], to solve (S1).

## Hard thresholding

Hard thresholding (HT) operator, [S5], can be applied entry-wise to **x** in (4) according to:  $b_{nt} = HT(x_{nt}) = \begin{cases} x_{nt} & \text{if } x_{nt} \ge \tau_1 \\ 0 & \text{if } x_{nt} < \tau_1 \end{cases}, n=1,...,N, t=1,...,T \text{ and } \tau_1 \in [10^{-6}, 10^{-4}] \text{ stands for a threshold.} \end{cases}$ 

HT preprocessing transform of  $\mathbf{x}$  yields vector  $\mathbf{b}$  with the same structure as  $\mathbf{a}$  given by eq.(6).

#### Soft thresholding

Soft thresholding (ST) operator, [S5], can be applied entry-wise to **x** in (4) according to  $c_{nt} = ST(x_{nt}) = \max(0, x_{nt} - \tau_2), n=1,...,N, t=1,...,T$  and  $\tau_2 \in [10^{-6}, 10^{-4}]$ . ST preprocessing transform of  $\mathbf{x}$  yields vector  $\mathbf{c}$  that, as  $\mathbf{b}$  obtained by HT, is also expected to have the same structure as  $\mathbf{a}$  in (6).

# Trimmed thresholding

Trimmed thresholding (TT) operator, [S6], is applied entry-wise to x in (4) according to:

$$d_{nt} = TT(x_{nt}) = \begin{cases} x_{nt} \frac{x_{nt}^{\alpha} - \tau_{3}^{\alpha}}{x_{nt}^{\alpha}} & \text{if } x_{nt} \ge \tau_{3} \\ 0 & \text{if } x_{nt} < \tau_{3} \end{cases} \text{ , } n=1,...,N \text{ , } t=1,...,T \text{ and } \tau_{3} \in [10^{-6}, 10^{-4}]. \alpha \text{ is a trade-}$$

off parameter between hard and soft thresholding. When  $\alpha=1$ , TT equals ST. When  $\alpha \rightarrow \infty$  TT is equivalent to HT. Herein, we set  $\alpha=3.5$  because this value yields TT to operate between ST and HT [S6]. TT preprocessing transform of **x** yields vector **d** that, as **b** obtained by HT and **c** obtained by ST, is also expected to have the same structure as **a** in (6).

#### Selection of threshold values

Threshold values suggested above can be justified by the following analysis. Due to A1 and A2 elements of  $\mathbf{g}^1$  and  $\mathbf{g}_1^{(2)}$  in (4) are less than 1. In pursuing worst case analysis of third-order effects we assume that third-order derivatives coefficients in  $\mathbf{g}_1^{(3)}$  are less than some value  $g_3$ . Thus, contribution of third-order terms is limited by above by  $x^{(3)} = M^{(3)}g_3s$ , where  $M^{(3)} = \begin{pmatrix} M+2\\ 3 \end{pmatrix}$  denotes number of 3rd order terms. If mixture value  $x_{nt}$  is greater than  $x^{(3)}$  then

it is probably due to first and second-order terms. The threshold value evidently depends on values of  $M^{(3)}$ ,  $g_3$  and s. For example, assuming M=100 ( $M^{(3)}=171700$ ),  $g_3=0.1$  and  $s=3.4\times10^{-7}$ 

we get  $x^{(3)}=5.8\times10^{-3}$ . However, that is overly pessimistic given the fact that most of the thirdorder cross-products will, due to sparseness assumption A3, vanish. Thus, optimal threshold value is somewhere in the interval  $[10^{-6}, 10^{-4}]$ .

# **References:**

S1. Candès E J, Li X, Ma Y, Wright H. Robust Principal Component Analysis? *J. ACM* 2011;58: Article 11 (37 pages).

S2. Chandrasekaran V, Sanghavi S, Paririlo P A, Wilsky A S. Rank-Sparsity Incoherence for Matrix Decomposition. *SIAM J. Opt.* 2011; **21**: 572-596.

S3. Lin Z, Ganesh A, Wright J, Wu L, Chen M, Ma Y. Fast Convex Optimization Algorithms for Exact Recovery of a Corrupted Low-Rank Matrix. *UIUC Technical Report UILU-ENG-09-2214*, August 2009.

S4. The website on low-rank matrix recovery and completion via convex optimization: http://perception.csl.illinois.edu/matrix-rank/sample\_code.html [22 May 2015]

S5. Donoho D L. De-Noising by Soft-Thresholding. *IEEE Trans. Inf. Theory* 1995; **41** (3): 613-627.

S6. Fang H T, Huang D S. Wavelet de-noising by means of trimmed thresholding. in: *Proc. of the 5th World Congress on Intelligent Control and Automation*, June 15-19, 2004, Hangzhou, P. R. China, pp. 1621-1624.

Description of chemical reactions: first reaction. L-Leucine (200 mg, 1.52 mmol) was dissolved in 5 mL of dry dimethylformamide (DMF) and solution was cooled to 0 °C. Nmethylmorpholine (NMM, 3.05 mmol, 337 µL) and isobutylchloroformate (IBCF, 3.34 mmol, 458  $\mu$ L) were added. Aliquots of the reaction mixture (100 $\mu$ L) were withdrawn every 30 minutes (t<sub>0</sub>-t<sub>8</sub>) solvent was evaporated and the residue dissolved in 1mL of 0.1 % formic acid (FA) in 50 % MeOH. Aliquots (100  $\mu$ L) were diluted with 400  $\mu$ L of 0.1 % FA in 50 % MeOH and 10  $\mu$ L were injected through autosampler on a column (Zorbax XDB C18, 3.5 µm, 4.7 mm) at the flow rate of 0.5 mL/min. Mobile phase was 0.1 % FA in water (solvent A) and 0.1 % FA in MeOH (solvent B). Gradient was applied as follows: 0 min 40 % B; 0-15 min 90 %B; 12-15min 90% B; 17.1 min 40% B; 17.1-20 min 40 %B. Electrospray ionization-mass spectrometry (ESI-MS) measurements operating in a positive ion mode were performed on a HPLC-MS triple quadrupole instrument equipped with an autosampler (Agilent Technologies, Palo Alto, CA, USA). The desolvation gas temperature was 300<sup>o</sup>C with flow rate of 8.0 L/min. The fragmentor voltage was 135 V and capillary voltage was 4.0 kV. Mass spectra were recorded in m/z segment of 10-2000. All data acquisition and processing was performed using Agilent MassHunter software. Acquired mass spectra are composed of intensities at T=9901 m/z coordinates.

Second reaction. *N*-acetylglycine (50 mg, 0.43 mmol), and Boc-Ala-OH (100 mg, 0.53 mmol) were dissolved in 5 mL of dry dimethylformamide (DMF) and solution was cooled to 0 °C. N-methylmorpholine (NMM, 3.62 mmol, 400  $\mu$ L) and isobutylchloroformate (IBCF, 3.65 mmol, 500  $\mu$ L) were added. After 5 minutes, L-valine (100 mg, 0.85 mmol), L-leucine (100 mg, 0.76 mmol) and L-phenylalanine (50 mg, 0.30 mmol) were added. Aliquots of the reaction mixture (100 $\mu$ L) were withdrawn every 15 minutes (t<sub>0</sub>-t<sub>11</sub>), solvent was evaporated and the residue

dissolved in 1mL of 0.1 % formic acid (FA) in 50 % MeOH. Aliquots (100  $\mu$ L) were diluted with 400  $\mu$ L of 0.1 % FA in 50 % MeOH and 10  $\mu$ L were injected through autosampler on a column (Zorbax XDB C18, 3.5 lm, 4.6975 mm) at the flow rate of 0.5 mL/min. Mobile phase was 0.1 % FA in water (solvent A) and 0.1 % FA in MeOH (solvent B). Gradient was applied as follows: 0 min 40 % B; 0-15 min 90 %B; 12-15min 90% B; 17.1 min 40% B; 17.1-20 min 40 % B. Electrospray ionization-mass spectrometry (ESI-MS) measurements operating in a positive ion mode were performed on a HPLC-MS triple quadrupole instrument equipped with an autosampler (Agilent Technologies, Palo Alto, CA, USA). The desolvation gas temperature was 300 °C with flow rate of 8.0 L/min. The fragmentor voltage was 135 V and capillary voltage was 4.0 kV. Mass spectra were recorded in *m*/*z* segment of 10-2000. All data acquisition and processing was performed using Agilent MassHunter software.

Third reaction. L-alanine (50 mg, 0.43 mmol), and L-Phe-OH (100 mg, 0.53 mmol) were dissolved in 5 mL of dry dimethylformamide (DMF) and solution was cooled to 0 °C. N-methylmorpholine (NMM, 3.62 mmol, 400  $\mu$ L) and isobutylchloroformate (IBCF, 3.65 mmol, 500  $\mu$ L) were added. After 5 minutes, L-proline (50 mg, 0.43 mmol) was added. Aliquots of the reaction mixture (100 $\mu$ L) were withdrawn every 10 minutes (t<sub>0</sub>-t<sub>11</sub>). L-Leucine (100 mg, 0.76 mmol) was added after 10 minutes, Boc-Gly-propargylamide (50 mg, 24 mmol) after 40 minutes and L-tyrosine (50 mg, 0.28 mmol) after 70 minutes. Solvent was evaporated and the residue dissolved in 1mL of 0.1 % formic acid (FA) in 50 % MeOH. Aliquots (100  $\mu$ L) were diluted with 400  $\mu$ L of 0.1 % FA in 50 % MeOH and 10  $\mu$ L were injected through autosampler on a column (Zorbax XDB C18, 3.5 lm, 4.6975 mm) at the flow rate of 0.5 mL/min. Mobile phase was 0.1 % FA in water (solvent A) and 0.1 % FA in MeOH (solvent B). Gradient was applied as

follows: 0 min 40 % B; 0-15 min 90 %B; 12-15min 90% B; 17.1 min 40% B; 17.1-20 min 40 % B. Electrospray ionization-mass spectrometry (ESI-MS) measurements operating in a positive ion mode were performed on a HPLC-MS triple quadrupole instrument equipped with an autosampler (Agilent Technologies, Palo Alto, CA, USA). The desolvation gas temperature was  $300^{\circ}$ C with flow rate of 8.0 L/min. The fragmentor voltage was 135 V and capillary voltage was 4.0 kV. Mass spectra were recorded in *m*/*z* segment of 10-2000. All data acquisition and processing was performed using Agilent MassHunter software. Acquired mass spectra are composed of intensities at *T*=9901 *m*/*z* coordinates.

**Mass spectrometry measurements.** Electrospray ionization-mass spectrometry (ESI-MS) measurements operating in a positive ion mode were performed on a HPLC-MS triple quadrupole instrument equipped with an autosampler (Agilent Technologies, Palo Alto, CA, USA). The desolvation gas temperature was  $300^{\circ}$ C with flow rate of 8.0 L/min. The fragmentor voltage was 135 V and capillary voltage was 4.0 kV. Mass spectra were recorded in *m/z* segment of 10-2000. All data acquisition and processing was performed using Agilent MassHunter software. Acquired mass spectra are composed of intensities at *T*=9901 *m/z* coordinates.



Figure S-1.

**Table S-1.** First nonlinear chemical reaction. Normalized cross-correlation coefficients between 25 pure components ( $s_1$  to  $s_{25}$ ) generated in nonlinear chemical reaction of peptide synthesis. Thereby, pairs of pure components are identified with normalized correlation coefficient above 0.1. Their mass spectra are shown in Figure S-3.

	<b>s</b> <sub>2</sub>	<b>s</b> <sub>6</sub>	<b>S</b> <sub>7</sub>	<b>S</b> 9	$\mathbf{s}_{10}$	<b>s</b> <sub>12</sub>
<b>S</b> <sub>1</sub>	0.9839	0.1416	0.1218	0.1796	0.1072	0.3343
	<b>s</b> <sub>6</sub>	<b>\$</b> 7	<b>S</b> 9	<b>s</b> <sub>10</sub>	<b>s</b> <sub>12</sub>	
<b>s</b> <sub>2</sub>	0.1418	0.1268	0.1797	0.1075	0.3305	
	<b>S</b> <sub>16</sub>	<b>S</b> 17	<b>S</b> <sub>18</sub>			
<b>S</b> <sub>3</sub>	0.3575	0.3103	0.1716			
	<b>s</b> <sub>6</sub>	<b>S</b> <sub>19</sub>	<b>s</b> <sub>21</sub>			
$\mathbf{S}_4$	0.3077	0.3947	0.4005			
	<b>S</b> <sub>7</sub>					
<b>S</b> <sub>5</sub>	0.7824					
	<b>S</b> 9					
<b>S</b> <sub>7</sub>	0.3297					
	<b>s</b> <sub>13</sub>					
$\mathbf{S}_8$	0.1293					
	<b>s</b> <sub>12</sub>	<b>S</b> <sub>22</sub>				
$\mathbf{s}_{11}$	0.2666	0.1622				
	<b>s</b> <sub>17</sub>					
<b>S</b> <sub>14</sub>	0.1024					
	<b>s</b> <sub>22</sub>					
<b>s</b> <sub>15</sub>	0.1349					
	<b>s</b> <sub>17</sub>					
<b>s</b> <sub>16</sub>	0.9783					
	<b>S</b> <sub>18</sub>					
$\mathbf{s}_{17}$	0.1186					
	<b>s</b> <sub>21</sub>					
<b>S</b> <sub>19</sub>	0.9962					
	<b>s</b> <sub>24</sub>	<b>S</b> <sub>25</sub>				
<b>S</b> <sub>23</sub>	0.4409	0.4339				
	<b>s</b> <sub>25</sub>					
<b>s</b> <sub>24</sub>	0.3008					











Figure S-2.











Figure S-3.





























Figure S-4.



























Figure S-5.



























Figure S-6.



Figure S-7.

**Table S-2.** Second nonlinear chemical reaction. Normalized cross-correlation coefficients between 19 pure components ( $s_1$  to  $s_{19}$ ) generated in nonlinear chemical reaction of peptide synthesis. Thereby, pairs of pure components are identified with normalized correlation coefficient above 0.1.

	<b>s</b> <sub>2</sub>	<b>S</b> <sub>3</sub>	<b>S</b> <sub>4</sub>	<b>\$</b> 5	<b>s</b> <sub>6</sub>	<b>S</b> 9	$\mathbf{s}_{10}$	$\mathbf{s}_{12}$	$\mathbf{s}_{14}$	<b>S</b> <sub>15</sub>	$\mathbf{S}_{18}$	<b>S</b> <sub>19</sub>
$\mathbf{s}_1$	0.9278	0.3376	0.6331	0.2617	0.3986	0.1417	0.1948	0.1103	0.3054	0.2102	0.2793	0.1899
	<b>\$</b> 3	<b>s</b> <sub>4</sub>	<b>S</b> 5	<b>s</b> <sub>6</sub>	<b>S</b> 9	<b>s</b> <sub>10</sub>	<b>s</b> <sub>12</sub>	<b>s</b> <sub>14</sub>	<b>s</b> <sub>15</sub>	<b>S</b> <sub>18</sub>	<b>S</b> <sub>19</sub>	
$\mathbf{s}_2$	0.3188	0.5896	0.2430	0.3717	0.1321	0.1815	0.11022	0.2870	0.1959	0.2599	0.1770	
	<b>S</b> 4	<b>S</b> <sub>6</sub>	<b>S</b> <sub>14</sub>	<b>S</b> <sub>18</sub>								
<b>S</b> <sub>3</sub>	0.2175	0.1550	0.1220	0.1082								
	<b>\$</b> 5	<b>s</b> <sub>6</sub>	<b>s</b> <sub>10</sub>	<b>s</b> <sub>14</sub>	<b>s</b> <sub>15</sub>	<b>s</b> <sub>18</sub>	<b>S</b> <sub>19</sub>					
$\mathbf{s}_4$	0.1764	0.2589	0.1266	0.2009	0.1506	0.2047	0.1788					
	<b>s</b> <sub>6</sub>	<b>s</b> <sub>15</sub>										
<b>S</b> 5	0.4043	0.2294										
	<b>s</b> <sub>13</sub>											
$\mathbf{s}_{11}$	0.5526											
	<b>s</b> <sub>14</sub>											
$\mathbf{s}_{12}$	0.3097											
	<b>S</b> <sub>15</sub>	<b>S</b> <sub>18</sub>										
$\mathbf{s}_{14}$	0.1284	0.1077										
	<b>S</b> <sub>18</sub>											
<b>S</b> 15	0.8788											
	<b>S</b> <sub>19</sub>											
$\mathbf{s}_{18}$	0.1185											



Figure S-8.

**Table S-3.** Third nonlinear chemical reaction. Normalized cross-correlation coefficients between 28 pure components ( $s_1$  to  $s_{28}$ ) generated in nonlinear chemical reaction of peptide synthesis. Thereby, pairs of pure components are identified with normalized correlation coefficient above 0.1.

	<b>s</b> <sub>2</sub>	<b>S</b> <sub>3</sub>	<b>s</b> <sub>16</sub>	<b>S</b> <sub>27</sub>		
$\mathbf{s}_1$	0.104	0.108	0.1932	0.1094		
	<b>S</b> <sub>3</sub>	<b>s</b> <sub>16</sub>	<b>s</b> <sub>18</sub>	<b>S</b> <sub>23</sub>	<b>s</b> <sub>24</sub>	<b>S</b> <sub>27</sub>
$\mathbf{s}_2$	0.9587	0.2014	0.1569	0.3261	0.4589	0.9099
	<b>S</b> <sub>16</sub>	<b>S</b> <sub>18</sub>	<b>S</b> <sub>23</sub>	<b>S</b> 24	<b>S</b> <sub>27</sub>	
<b>S</b> <sub>3</sub>	0.1883	0.1553	0.3171	0.4426	0.8807	
	<b>s</b> <sub>21</sub>					
<b>s</b> <sub>6</sub>	0.7299					
	<b>s</b> <sub>12</sub>	<b>s</b> <sub>15</sub>				
$\mathbf{s}_{10}$	0.1094	0.2294				
	<b>s</b> <sub>13</sub>					
$\mathbf{s}_{11}$	0.1229					
	<b>s</b> <sub>16</sub>	<b>S</b> <sub>17</sub>	<b>S</b> <sub>28</sub>			
$\mathbf{s}_{15}$	0.1128	0.1440	0.1171			
	<b>S</b> 27					
$\mathbf{s}_{16}$	0.1913					
	<b>s</b> <sub>27</sub>					
$\mathbf{S}_{18}$	0.1591					
	<b>s</b> <sub>21</sub>	<b>s</b> <sub>23</sub>				
$\mathbf{s}_{20}$	0.1534	0.3122				
	<b>s</b> <sub>23</sub>					
$\mathbf{s}_{21}$	0.2583					
	<b>s</b> <sub>24</sub>	<b>s</b> <sub>27</sub>				
<b>S</b> <sub>23</sub>	0.1563	0.3109				
	<b>s</b> <sub>27</sub>					
<b>s</b> <sub>24</sub>	0.4561					
	<b>S</b> <sub>28</sub>					
$\mathbf{s}_{26}$	0.8870					