

## Determining the Number of Components for Multivariate Curve Resolution: Case Study Using Raman Mapping of Pharmaceutical Tablets

Clémence Fauteux-Lefebvre<sup>1</sup>, Francis B. Lavoie<sup>1</sup>, Marie-Josée Colbert<sup>1</sup>, Jean-Maxime Guay<sup>2</sup> and Ryan Gosselin<sup>1</sup>

<sup>1</sup>. Université de Sherbrooke, Department of Chemical and Biotechnological Engineering, Sherbrooke, Canada.

<sup>2</sup>. Process Analytical Sciences Group, Pfizer Global Supply, Montréal, Canada.

Raman chemical imaging of pharmaceutical tablets can be used to determine particle morphology, active pharmaceutical ingredient (API) concentration, sample heterogeneity, as well as for impurity detection. Raman imaging generally requires simple and fast sample preparation and offers  $\mu\text{m}$ -scale spatial resolutions. The probe used in this study (mPAT™ Raman system from H<sub>2</sub>Opx) also performs 3D analysis of solid samples. In pharmaceutical tablet composition analysis, there is an increased interest for Raman mapping using chemometric tools such as multivariate curve resolution (MCR) [1].

Physically meaningful spectra can be extracted using some MCR variants, which is especially useful when dealing with unknown compound. However, one of the difficulties with MCR is fixing the number of components [1-2]. The raw spectra depend on the acquisition parameters, chemical nature and number of compounds in the sample as well as their distribution, via the spatial resolution and point spread function of the probe. The computed MCR spectra depend both upon the raw spectra and the number of user-defined components. This point is critical when the percentage of mixed phases (pixels with more than one component) is high, a typical case for pharmaceutical tablet with macroscopic homogeneity and wide range of particle sizes.

This work focuses on the impact of the number of MCR components and the particle size distribution on the results of the Raman analysis using model tablets. Two tablet types were produced: type 1 only contains coarse particles (greater than  $90\mu\text{m}$ ) whereas type 2 contains a blend of coarse and fine particles (smaller than  $90\mu\text{m}$ ). Powders with different and easily distinguishable Raman signature were used: lactose, hydrate of zinc sulfate and trisodium citrate.

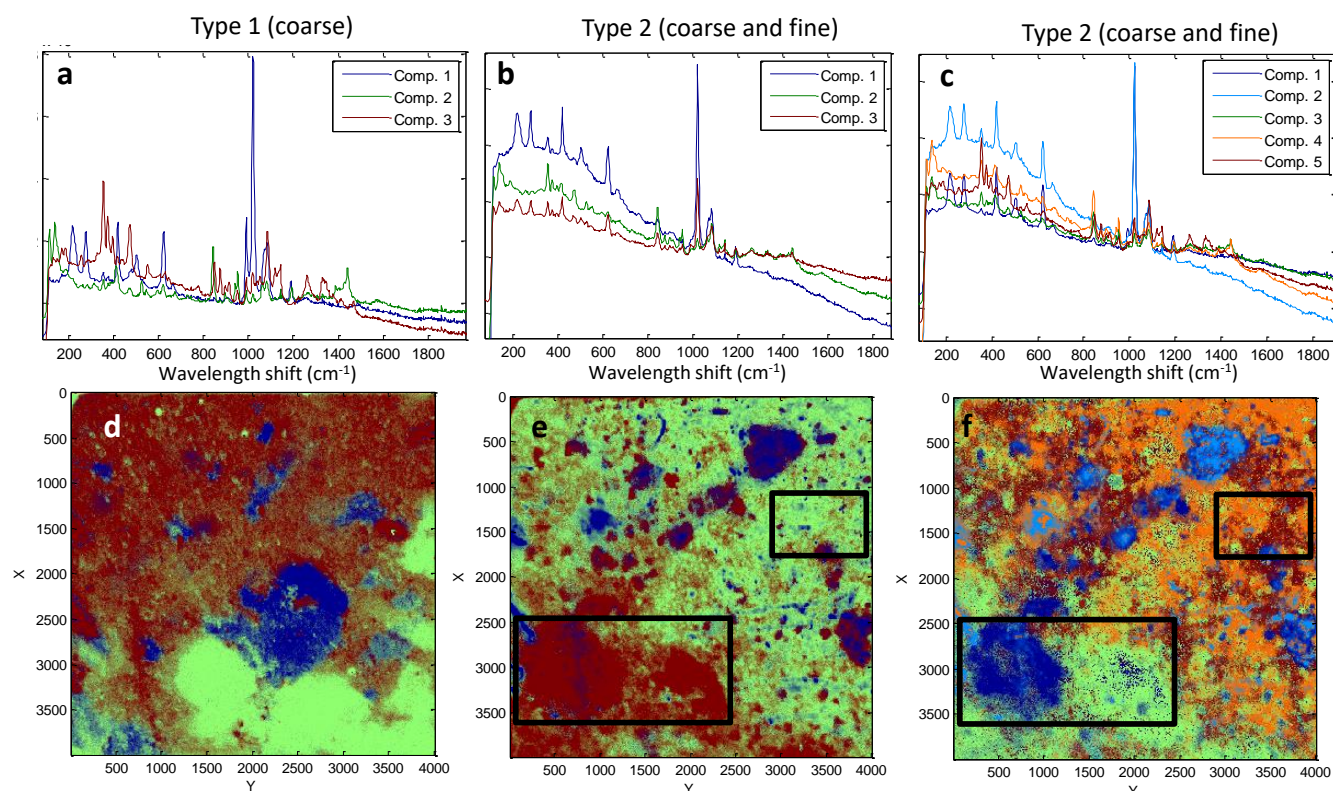
A  $400 \times 400$  pixel Raman map was acquired for each sample, with a  $10\text{-}\mu\text{m}$  step size. The samples were illuminated by a  $785\text{ nm}$  laser, with an approximate spot size of  $5\mu\text{m}$ . The data generated were analyzed using MCR combined with Log-Likelihood Maximization (MCR-LLM), a method developed previously in the research group [3]. This method takes into account the noise characteristic of the signal and enhances the coherence of the extracted spectra, even with low count signals.

The extracted spectra are shown in figure 1 (a, b and c), for tablet types 1 and 2 (ternary blends). In figure 1a, 3 components were extracted for type 1 tablets (coarse) while in figure 1b, 3 components were extracted for type 2 tablets (coarse and fine). In type 1 tablets, the extracted spectra can be easily attributed: component 1 (blue) to zinc sulfate, component 2 (green) to trisodium citrate and component 3 (red) to lactose. As seen in figure 1d, the different phases in the corresponding map can be easily identified. On the other hand, the Raman signatures are confounded in type 2 tablets when only 3 components are extracted (figure 1b), with similar effect on the map, with inadequate segmentation and missing information (figure 1e).

Increasing the number of components improves the Raman signatures and composition maps as shown in figure 1c and 1f (refer to highlighted areas). However, the extracted spectra do not correspond to the spectra of the pure ingredients, but to their blends. Working with small particle sizes, the Raman map shows few, if any, pure pixels representing unblended raw materials. While the MCR algorithm can compute the “purest” forms present in the image, it cannot provide true raw material composition without human input (i.e. cross reference with known raw material spectra). Further increase in the number of components decrease the clarity for the interpretation of the maps but tends to improve spectral identification [4].

#### References:

- [1] A Farkas *et al*, Journal of Chemometrics **31** (2017), p. 1.  
 [2] H Noothalapati, K Iwasaki and T Yamamoto, Analytical Sciences **33** (2017), p. 15.  
 [3] F B Lavoie, N Braidy and R Gosselin, Chemometrics and Intelligent Laboratory Systems **153** (2016), p. 40.  
 [4] The authors acknowledge funding from Pfizer, NSERC and Mitacs.



**Figure 1.** Spectra extracted using MCR-LLM for ternary tablets (a,b,c) and concentration mapping (d,e,f). a) and d) Type 1: Coarse particles, 3 comp. b) and e) Type 2 : Coarse and fine particles, 3 comp. c) and f) Type 2: Coarse and fine particles, 5 comp. Highlighted areas are representative of the effect of the number of components.