1. INTRODUCTION

In comments published by Tauler [1] on the paper recently published by the present author titled ‘Some surprising properties of multivariate curve resolution alternating least squares (MCR-ALS) algorithms’ [2], Tauler concluded that ‘the discrepancies and surprising properties observed by Rajko in his paper were mostly due to machine accuracy errors and due to a too loose interpretation of rotation ambiguities in MCR-ALS results.’ In this rejoinder it is explained again that the revealed discrepancy of MCR-ALS algorithms (i.e. the sub- and even the final solutions can be outside the range of the data matrix), does exist and this theoretical fact could not be refuted by Tauler.

2. DEFINITIONS OF NEW TERMINOLOGY

2.1. Orthogonal part of estimations

Estimations of \( C \) or \( S \) can have orthogonal parts to the sub-space spanned by the range of \( R: C_{\text{MCR-ALS}}^{\text{Orth}} = (I - RR^T)C_{\text{MCR-ALS}}, \)
\( S_{\text{MCR-ALS}}^{\text{Orth}} = (I - R^T R)S_{\text{MCR-ALS}}. \)

2.2. Embedded part of estimations

Estimations of \( C \) or \( S \) should have embedded parts in the sub-space spanned by the range of \( R: C_{\text{MCR-ALS}}^{\text{Emb}} = (RR^T)C_{\text{MCR-ALS}}, \)
\( S_{\text{MCR-ALS}}^{\text{Emb}} = (R^T R)S_{\text{MCR-ALS}}. \)

2.3. Maxima of absolute valued orthogonal parts (MAVOP)

In all iterations of MCR-ALS maximum of \( \text{abs}(C_{\text{MCR-ALS}}^{\text{Orth}}) \) and \( \text{abs}(S_{\text{MCR-ALS}}^{\text{Orth}}) \) can be calculated and plotted.

2.4. Constraint of ‘zero orthogonal part for the estimated profile’

The embedded part of the estimated profiles should be used instead of the non-negative estimations lying outside the range of \( R \). More details can be found in Reference [2].

3. RESULTS AND DISCUSSION

Tauler [1] showed the triviality that increasing the number of iterations for MCR-ALS the final solution can be dramatically improved. For the first example containing simple structured data, Tauler demonstrated that ‘7138 iterations were needed giving a lack of fit of \( 2.7 \times 10^{-12}/12\% \). This lasted approximately 2–3 min (depends on computer power). The orthogonal parts of the solutions were extremely small \( (<10^{-13}) \) and also very small negative embedded parts \( (<-10^{-13}) \) were calculated.’ However, these results cannot refute the fact that the sub-solutions were outside the range of the data matrix \( R \) during the iteration; see the insets of Figure 2 in Reference [2], where all the iteration tracks are outside the outer polygon (OP) which is the boundary for the non-negative profiles. In addition, the left panel of Figure 1 shows the MAVOP of the concentration estimations of the three components during the iteration. It is distinctly visible that...
Figure 1. MAVOP of the concentration estimations of the three components during the iteration. Left panel: simple simulated data. Right panel: complex simulated data.

MAVOP start relatively large values and they decrease exponentially with the iteration steps (their logarithmic transformations are linear after the initial part) and they can be negligible only after 6000 iterations.

For the second example containing complex structured data, Tauler showed that ‘the convergence (relative change of the standard deviations of the residuals <0.1%) was achieved in 628 iterations (approximately 1 min) with a lack of fit of $6.2 \times 10^{-05}$.

In this case also, no negative parts were obtained in the profiles calculated by MCR-ALS. The maximum of the embedded concentration peak of the second component (the more difficult to resolve) was 0.1514 and the orthogonal part of the same peak was $9.327 \times 10^{-05}$. This was only $0.0616\%$!

First of all, using my constraint ‘zero orthogonal part for the estimated profiles’ the following result can be obtained: $\sim 10^{-12}\%$ lack of fit in 53 iterations. The benefit of using the suggested constraint is obvious (cf. Figures SM-33–SM-38 in Supporting Information of Reference [2] as well for seeing better orthogonal and embedded parts, i.e. less (closer to zero) orthogonal and greater (farther from zero) embedded parts). On the other hand, the MCR-ALS estimated that the value of the second component is zero at the 75th concentration position (i.e. at 75th virtual

Figure 2. Borgen plot of concentration space for the complex simulated data with iteration tracks provided by MCR-ALS in 628 iterations. The OP (larger red polygon) is the inner ‘hole’ of the blue lines which define positive and negative valued half-spaces (inside of the ‘hole’, i.e. the OP, the points represent retransformed profiles having only non-negative values). The smaller red polygon is the inner polygon calculated by using the transformed data points. The black dashed-dotted lines form the two Borgen simplexes (here triangles). The feasible regions (blue shaded areas) are in the area between the outer and inner polygons. The vertices of the green dashed triangle mark the true concentrations. The small magenta circles represent the iteration tracks. The vertices of the magenta triangle mark the final estimated concentrations calculated by MCR-ALS.
elution time). At this point, the orthogonal part of the estimation got its maximum. Thus, the value of $\sim 10^{-4}$ is non-negligible comparing it to zero (the comparison should not be made to the maximum of the estimated values). Figure 2 shows the Borgen plot of concentration space for the complex-simulated data with iteration tracks provided by MCR-ALS in 628 iterations. The estimation of the second component is definitely outside of and very far from the OP, meaning that its orthogonal part is non-negligible. The right panel of Figure 1 depicts MAVOP of the concentration estimations of all three components during the iteration. Because of the noiseless case, all MAVOP are significantly larger than zero.

Figures SM-2–SM-5 in the Supporting Information show that using reasonably chosen initial guesses, even the final solution can be outside the range of $R$ and the feasible regions after exhaustive iterations. In Figure SM-2, the iteration track of the concentration estimation of component 2 approaches the feasible region at the beginning, but after that it recedes from the OP.

In Figure SM-4, the iteration track of the spectrum estimation of component 2 is first inside the OP. However, after that the track steps out of the OP and it approaches the feasible region, but it does not reach the OP again in 12 067 iterations.

4. CONCLUSIONS

The main conclusion may be that the theory and the practice should not be confused.

Theoretically, the revealed discrepancy exists; however, Tauler is right that the effect of this discrepancy can be irrelevant for some practical situations. The price is the longer convergence time (see the perfect results after 10 000 iterations in Figures SM-31 and SM-32 in the Supporting Information of Reference [2]). However, for real-time applications of MCR-ALS the convergence time becomes crucial. Using the newly introduced constraint ‘zero orthogonal part for the estimated profiles’ the number of iterations and the convergence time were reduced and the quality of the solution improved as mentioned above.

Moreover, the relatively high amount of the orthogonal parts of the estimations can indicate the non-perfect convergence. In that case, the user can apply more restrictive convergence criteria until the orthogonal parts will be close to zero.

The final question is the noise. I showed my results using noiseless data [2], because my investigation was theoretical-like. For analytical measurements, the signal-to-noise ratio (SNR) should be increased to improve the quality of the obtained information. It can be done (1) to develop new and much more selective, sensitive, etc. analytical instruments/procedures, and/or (2) to use noise reduction (even elimination) signal processing methods [3]. Due to the latter possibility, chemometric methods which need noiseless data (like the Borgen plot method) could become relevant even in practice (and not just in theory) now or in the near future.

REFERENCES