Inline Raman Spectroscopy and Indirect Hard Modeling for Concentration Monitoring of Dissociated Acid Species

Alexander Echtermeyer[1], Caroline Marks[1], Alexander Mitsos[3,1,2] and Jörn Viell[1,*]

[1] Process Systems Engineering (AVT.SVT), RWTH Aachen University, 52074 Aachen, Germany

[2] Institute for Energy and Climate Research IEK-10: Energy Systems Engineering, Forschungszentrum Jülich, 52425 Jülich, Germany

[3] JARA-ENERGY, Templergraben 55, 52056 Aachen, Germany

Corresponding author: *J. Viell
Process Systems Engineering (AVT.SVT), RWTH Aachen University, Forckenbeckstr. 51, 52074 Aachen, Germany
E-mail: joern.viell@rwth-aachen.de

Abstract

We propose an approach for monitoring the concentration of dissociated carboxylic acid species in dilute aqueous solution. The dissociated acid species are quantified employing inline Raman spectroscopy in combination with Indirect Hard Modeling (IHM) and Multivariate Curve Resolution (MCR). We introduce two different titration-based Hard Model (HM) calibration procedures for a single mono- or polyprotic acid in water with well-known (method A) or unknown (method B) acid dissociation constants pK_a. In both methods, spectra of only one acid species in water are prepared for each acid species. These spectra are used for the construction of HMs. For method A, the HMs are calibrated with calculated ideal dissociation equilibria. For method B, we estimate pK_a values by fitting ideal acid dissociation equilibria to acid peak areas that are obtained from a spectral HM. The HM in turn is constructed on the basis of MCR data. Thus, method B on the basis of IHM is independent of a priori known pK_a values, but instead provides them as part of the calibration procedure. As a detailed example, we analyze itaconic acid in aqueous solution. For all acid species and water, we obtain low HM errors of less than 2.87×10^{-4} mol mol^{-1} in the cases of both method A and B. With only four calibration samples, IHM yields more accurate results than partial least squares regression. Furthermore, we apply our approach to formic, acetic, and citric acid in water, thereby verifying its generalizability as a process analytical technology for quantitative monitoring of processes containing carboxylic acids.

Keywords

Process analytics; Indirect Hard Modeling; IHM; Raman spectroscopy; Multivariate Curve Resolution; MCR-ALS; Itaconic acid; Carboxylic acid; Dissociation; pK_a estimation

1 Introduction

Carboxylic acids are predicted to play a major role as intermediate products or versatile platform chemicals in a biobased economy.[1–10] The key factor in achieving a cost-effective, competitive process is the interplay of production and separation of the carboxylic acids.[11,12] The efficiency of integrated separation and purification strategies such as crystallization and extraction is often determined by dissociation of the acids as, for example, only the associated form of the acids is extracted.[13–16] Therefore, robust and time-resolved analytics to measure the concentration and dissociation of the different acid species is necessary for modeling, control, and optimization of downstream process unit operations.[17]

The performance of such unit operations is usually quantified by concentration measurements. In aqueous solution, accurate knowledge of the corresponding acid dissociation constant pK_a, pH and total acid concentration is sufficient to calculate the concentrations of acid species. This does not hold for mixtures with multiple chemical species influencing the pH value or dissociation of acids at high concentrations and high ionic strengths. These conditions typically occur in acid purification processes. Hence, accurate concentration data is the key to monitoring acid production and purification processes.

In general, acid concentrations can be provided by offline or atline analytics such as gas and high-performance liquid chromatography.[2,4,6,8,13,14,18,19] These methods are widely used, but they require process sampling and are time-consuming. Moreover, they usually fail to distinguish between different acid dissociation stages inline.

Inline spectroscopy can overcome the mentioned sampling difficulties as it enables direct measurements of chemical species in the process at short acquisition times usually in the range of a few seconds.[18,20,21] Infrared (IR) and Raman spectroscopy resolve fundamental molecular vibrations and therefore enable the identification of different chemical species in a mixture. The latter technique is especially suited for process monitoring of dissociated acid species in aqueous solution. The reason for this is the rather low Raman activity of water, which increases sensitivity for dissolved acids.

The use of Raman or complementary IR spectroscopy for quantitative process monitoring of acid species is rather scarce in the literature. In most cases, the dissociation of strong mineral acids is investigated for fundamental research on species interaction and thermodynamic parameters.[22–26] H-bond formation, hydration number, and dissociation of carboxylic acids have been investigated by factor analysis and IR spectroscopy in combination with titration and selective water spectra subtraction for quantitative calibration.[27–29] The approach has been suggested for Raman spectroscopy but has not yet been realized.[29] Some studies approach dissociation of strong mineral acids with Raman spectroscopy to develop process analytical tools for determination of the acid dissociation ratio.[30,31] However, they either rely on available pK_a values or use an empirical approach for quantitative evaluation. Besides analysis of mineral acids, inline process monitoring of carboxylic acid species by Raman and IR spectroscopy has been successfully demonstrated during crystallization for liquid and solid phases, for characterization of citric acid and salicylic acid in aqueous solution, and for dissociation of gallic acid in aqueous solution.[32–36] Some approaches have already succeeded in determining unknown pK_a values.[34–36] However, despite of a few studies, estimation of pK_a was not part of process analytical technology so far.

Quantitative evaluation of Raman spectra requires a chemometric method to correlate the signal information with the concentration of acid species. The spectroscopic signal at each wavenumber relates to the number of molecules within the spectroscopic measurement volume. This signal results in a spectrum whose intensities of peaks is proportional to concentration (cf. Beer-Lambert law).[37] In case of overlapping peaks or small molecular variations, i.e., associated versus dissociated acid species, more sophisticated methods have to be used for quantification. A large variety of chemometric methods are available, of which peak integration (PI) and partial least squares (PLS) regression are the most popular.[20,21,38–41] PI calculates the area of a peak which is proportional to the concentration for the respective chemical species.[37] However, PI requires non-overlapping peaks. Such peaks are scarce for structurally similar components such as dissociated acid species with similar spectral signals. PLS regresses calibration spectra on corresponding concentration data to find the maximum covariance in the two data sets. This

is achieved by decomposition of the data matrix into orthogonal principal components sorted by descending eigenvalues.[38,39,42] PLS enables a fast evaluation of complex overlapping mixture spectra, but as a bilinear calibration model, it is limited in assessing non-linear peak shifts and peak deformations caused by molecular interactions in electrolytic media. This could be overcome by non-linear PLS or locally weighted regression, for example by multiple local PLS models, however at the expense of additional complexity and computational cost.[43,44] Furthermore, PLS is purely data-driven. It therefore relies on a rather large set of calibration data and lacks the physical depth of the calibration compared to IHM.

Indirect Hard Modeling (IHM) overcomes the drawbacks of PI and PLS by using linear combinations of Gaussian and Lorentzian functions (i.e., pseudo-Voigt profiles) to form spectral Hard Models (HMs) of each chemical species contributing to the selected spectral range.[45–47] These physically justified Pure Component Models (PCMs) are integrated, weighted and superimposed to construct a mixture HM. The latter can be subsequently regressed to mixture spectra by adjustment of PCM weights and a set of flexible peak parameters. Therefore, the PCM can be selective and can account for peak shifts and peak deformations in contrast to linear models. Moreover, a closure constraint that supports physically meaningful results can be considered for IHM in contrast to PLS calibration. Until now, inline Raman spectroscopy with IHM has not been reported for the quantitative monitoring of acid dissociation.

In the case of dissociated acid species, each species corresponds to one PCM. The construction of physically justified PCMs requires a systematic approach such as Complemental Hard Modeling (CHM), Hard Model Factor Analysis (HMFA), or Multivariate Curve Resolution (MCR).[47–51] In CHM, one or a set of PCMs is used to support the construction of another PCM on the basis of a mixture spectrum of k chemical species. However, CHM is only applicable for mixtures where PCMs of $k - 1$ chemical species are known. In HMFA, multiple PCMs are automatically obtained from a set of mixture spectra without the need for a priori known PCMs. Nevertheless, the algorithm requires the overall number of peaks to be modeled as input. In the case of chemical species with strong structural similarities, similar spectral signals impede

a priori identification of the correct number of peaks, therefore rendering this method inappropriate. Alternatively, MCR with Alternating Least Squares algorithm (MCR-ALS) can be used to directly compute pure component spectra from a set of mixture spectra. MCR-ALS forms a linear additive model subject to physical constraints such as non-negativity of concentrations and spectra, unimodal concentration profiles, and closure constraint, to describe the spectral data set as summed products of concentrations of species and their pure component spectra. As a main difference to PLS, MCR-ALS aims at providing physically meaningful pure component spectra and concentration profiles rather than an orthogonal decomposition of the spectral data in the form of abstract mathematical factors. However, MCR-ALS is a purely linear model and, similarly to PLS, not capable of assessing non-linear spectral effects. The combination of MCR-ALS with CHM for processing of unknown multi-component spectra to provide spectral data for IHM of PCMs, though, can overcome the disadvantages of the two single approaches.

In this work, we employ inline Raman spectroscopy combined with IHM to monitor the concentration and dissociation of carboxylic acids in aqueous solution. Titration and selective spectra subtraction enable the analysis of the two cases of known (method A) and unknown (method B) pK_a values. As a novelty, we present a combination of MCR-ALS, IHM and parameter estimation for the quantification of dissociated carboxylic acid species in the case of unknown pK_a values (method B). We test our methods on aqueous solution of itaconic, formic, acetic, and citric acid, respectively, and compare the calibration results of the IHM to those of PLS.

2 Materials and Methods

In the subsequent subsection, the overall approach for method A and B is outlined by the main steps and corresponding assumptions. This is followed by a more detailed view on the experimental procedure and the computational approach in the next subsections. Method A and B are explained in detail for diprotic itaconic acid but are shown to hold also for other carboxylic acids, such as monoprotic formic and acetic, as well as triprotic citric acid.

2.1 Overview of Method A and B

The four major tasks and corresponding steps for construction of a mixture HM for quantitative evaluation of the concentrations of each dissociation stage of one acid in aqueous solution for the two cases of known (method A) and unknown (method B) acid dissociation constants pK_a are schematically summarized in Fig. 1 and briefly outlined in the following. All calculations are based on the assumption of dilute, ideal mixtures, thus with activity coefficients equal to unity. We neglect changes of both the water concentration and excess volume during titration. Moreover, we do not differentiate between Raman signals of water, oxonium, and hydroxide ions but assign all three components to the overall water signal. We assume equal Raman activities for each acid dissociation stage.[52]

In the cases of both known and unknown pK_a, the first major task is the generation of two-species spectra that form the basis for PCM construction. In the case of known pK_a, the pH at which each individual acid species has its concentration maximum is calculated a priori. In the case of unknown pK_a, a priori concentration calculation is not yet possible at this stage. Next, Raman spectra of a titration experiment are recorded that is either a minimal set of titration samples with a priori calculated pH-based concentration data in the case of known pK_a or a titration of the full pH range in the case of unknown pK_a, respectively. The recorded Raman spectra are pretreated to eliminate external influences of the experimental setup, measurement device and environment. After pretreatment, the spectra are processed either by linearly weighted spectra subtraction in the case of known pK_a or MCR-ALS in the case of unknown pK_a to obtain two-species spectra containing signals of only one acid species and water.

The second major task is the construction of mixture HMs. The two-species spectra are used in IHM to construct PCMs by CHM for each dissociation stage of the acid and water. The PCMs are subsequently combined to mixture HMs.

The third major task of composition calculation via estimated pK_a values only applies to the case of unknown pK_a. Here, the mixture HM is first fitted to the titration spectra to obtain the corresponding ratios of peak areas of the acid species. Next, the dissociation equilibria are fitted

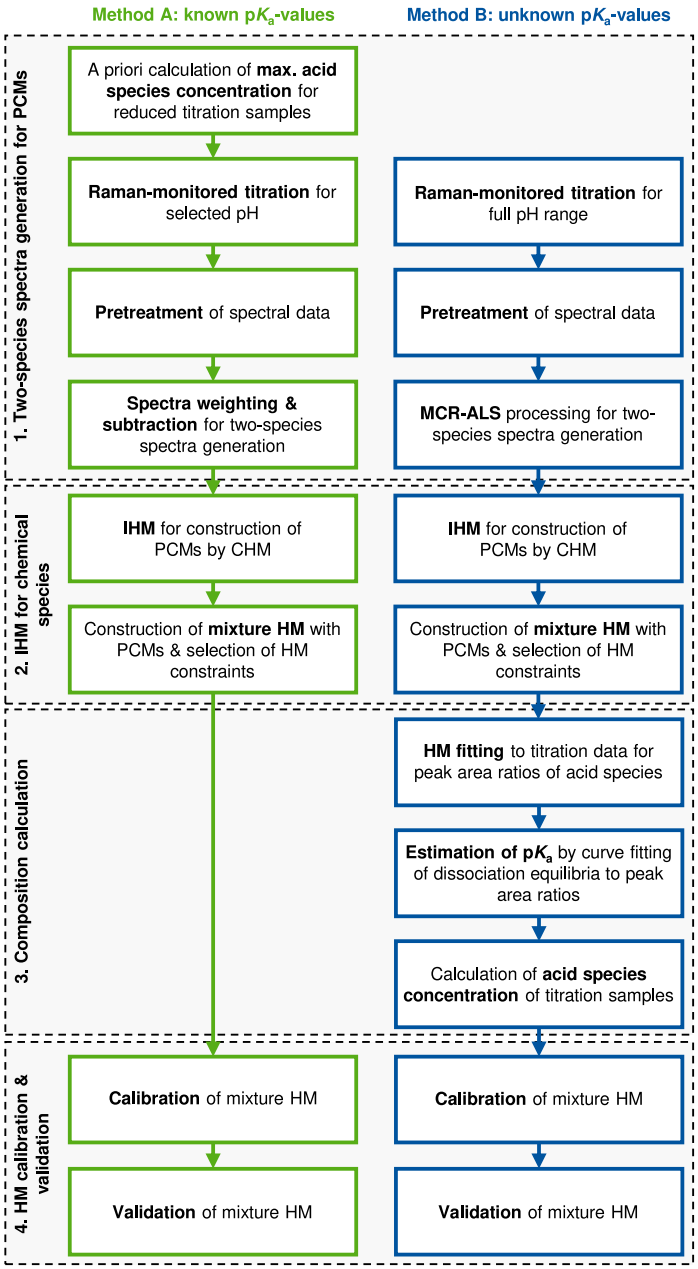

Figure 1. Overview of the four main tasks and corresponding steps to construct a mixture HM for quantitative monitoring of the concentrations of acid species in aqueous solution for the two cases of known (method A) and unknown (method B) acid dissociation constants pK_a.

to the ratios of peak areas of the acid species in an optimization approach to estimate the unknown pK_a values. Estimated pK_a values enable calculation of the titration sample composition in the following step.

The fourth and final major task is the mixture HM calibration that, in the cases of both known and unknown pK_a, is done on the basis of the titration Raman spectra and corresponding composition data.

2.2 Experimental

For titration, analyte solutions of itaconic acid (Alfa Aesar, $\geqslant$ 99%), formic acid (Merck, 98 – 100%), acetic acid (Merck, 99 – 100%), and citric acid (KMF Laborchemie, $\geqslant$ 99%) are prepared with defined concentrations in deionized water (conductivity 0.8 μScm^{-1} at 25 °C). A scheme of the experimental setup for titration is provided in Fig. S1, Supplemental Material (SM). Aqueous NaOH (VWR Chemicals, 1 mol L^{-1}) is used for titration of the analyte solution in a three-neck round bottom flask. The flask is equipped with a magnetic stirrer (250 rpm) and held at a constant temperature of 25 °C (T_{bath} in Fig. S1) in a stirred water bath. For exclusion of ambient light, the setup is covered with non-transparent black PVC liner. Temperature and pH of the analyte are measured continuously during titration using a SenTix 940 pH electrode with Multi 3420 pH meter (WTW). Inline Raman spectra are recorded for the analyte solution before titration and at equilibrium with constant pH during titration. Raman spectra are collected employing a RXN2 Raman analyzer with 400 mW at 785 nm (Kaiser Optical Systems).[53] The analyzer is equipped with fiber-optic cables and NIR immersion probes. Inline Raman spectra are acquired in a spectral range of 160 – 3285 cm^{-1} with 4 cm^{-1} resolution and 60 s acquisition time in HoloGRAMS ver. 3.2. Additionally, Raman spectra of pure deionized water are collected. Further information on Raman spectra collection and pH measurement is provided in the SM.

2.3 Modeling

2.3.1 Pretreatment of Spectral Data

Pretreatment of Raman spectra is done either in MATLAB ver. R2017a (The MathWorks, Inc.) for spectra subtraction or in PEAXACT ver. 4.5 (S-PACT) for all other steps.[46,54,55] GSTools ver. 0.4.2 is used for importing and exporting Raman data files to MATLAB, whereas mdatools ver. 0.1.6 is employed for spectra normalization by standard normal variate (SNV) method prior to spectra subtraction.[57,58] For all steps, a spectral range of 1025 – 1850 cm^{-1} is selected for evaluation with exclusion of the atmospheric oxygen signal between 1545 – 1565 cm^{-1} originating from air in the pathway of the laser beam in the Raman immersion optical probe.[53,59]

2.3.2 Subtraction of Raman Spectra

In the case of known pK_a (reported values for pK_{a1} = 3.85 and pK_{a2} = 5.45), a two-species spectrum of the first dissociation stage of itaconic acid (IA^-) in water is generated by spectra subtraction in MATLAB ver. R2017a.[54,60] Pretreated Raman spectra of fully protonated (IA) and fully de-protonated (IA^{2-}) itaconic acid in water are subtracted from a spectrum containing all three acids species in water (cf. SM for equations and further details).

2.3.3 Multivariate Curve Resolution

In the case of unknown pK_a, physically meaningful two-species spectra of each acid species in water are generated by MCR-ALS in PEAXACT.[48,50] The Raman spectra of titration samples with unknown concentration of the dissociated acid species are used as input for the MCR-ALS algorithm. The algorithm is restricted by four constraints: (1) non-negativity for concentrations, (2) non-negativity for spectra, (3) concentration profiles with a single maximum (unimodality), and (4) sum of concentrations equal to unity (closure constraint). A maximum number of 100 iterations and a default convergence tolerance of 1×10^{-5} is set. Offset subtraction for baseline correction is used.

2.3.4 Indirect Hard Modeling

For the construction of PCMs by IHM and CHM, as well as for the construction of mixture HMs, PEAXACT is used. A PCM of water is constructed by IHM via fitting of 4 pseudo-Voigt profiles to a Raman spectrum of pure deionized water. The resulting PCM of water is subsequently used for CHM to construct PCMs for each acid species.[47] IA is modeled by 9 pseudo-Voigt profiles, whereas for IA^- and IA^{2-}, 10 pseudo-Voigt profiles are needed to model the respective Raman signals (see Tab. SIV for modeling details on formic, acetic, and citric acid). All four PCMs are combined to a mixture HM with a linear baseline. A linear fit baseline correction and in the case of a priori known pK_a also SNV correction are used as additional pretreatment settings in the HM. A total number of 43 free parameters (4 component weights, 4 component shifts, 2 baseline parameters, and 1 peak position per pseudo-Voigt profile) of the mixture HM are automatically adjusted in PEAXACT for a spectral fit such that the mixture HM fits the measured Raman spectra well. A fixed component shift for IA^{2-} is set in the mixture HMs to avoid unphysical fitting. Otherwise, default settings for HM peak constraints are selected: component shift ± 10 cm^{-1}, position ± 30 cm^{-1}, maximum $\pm 30\%$, half width at half maximum (HWHM) $\pm 30\%$, and Gaussian part ± 0.5. However, the parameters for maximum, HWHM, and Gaussian part are fixed for the constructed HMs and therefore do not change from their initial value during HM fitting.

2.3.5 Dissociation Equilibria and Estimation of Acid Dissociation Constants

In the case of unknown values for pK_{a1} and pK_{a2} (defined in SM, Eq. (S5)-(S6)), the dissociation constants are estimated by minimizing the *RMSE* of experimentally determined ratios $x_{k,j,\mathrm{exp}}$ and calculated ratios $x_{k,j,\mathrm{calc}}$ of the acid species k for each titration step j out of the total number of titration steps J, $k \in [\mathrm{IA}, \mathrm{IA}^-, \mathrm{IA}^{2-}]$. This is valid by the assumption of equal Raman activities for the k acid species. Thereby, $x_{k,j,\mathrm{exp}}$ is defined as the peak area of the PCM of one acid species divided by the sum of the peak areas of the PCMs of all acid species k, which are obtained from fitting a mixture HM to the spectrum of each titration step j. Values for $x_{k,j,\mathrm{calc}}$ are

determined from the calibration sample composition on the basis of the acid species balance and acid dissociation equilibria containing the free parameters pK_{a1} and pK_{a2} (underlying equations for $f(pK_{a1}, pK_{a2})$ are provided in the SM, Eq. (S2)-(S15)). For better scaling of the optimization problem, we do not directly estimate K_{a1} and K_{a2} but rather use the negative decadic logarithm yielding pK_{a1} and pK_{a2} as degrees of freedom (DOFs). The optimization problem reads as follows:

$$\begin{aligned} \min_{pK_{a1},pK_{a2}} \quad & \sum_k RMSE_k \qquad \forall k \in [\mathrm{IA}, \mathrm{IA}^-, \mathrm{IA}^{2-}] \\ \text{s.t.} \quad & RMSE_k = \sqrt{\frac{1}{J}\sum_{j=1}^{J}\left(x_{k,j,\mathrm{exp}} - x_{k,j,\mathrm{calc}}\right)^2} \\ & x_{k,j,\mathrm{calc}} = f(pK_{a1}, pK_{a2}) \\ & 0 \leqslant pK_{a1} \leqslant 14 \\ & 0 \leqslant pK_{a2} \leqslant 14 \end{aligned} \tag{1}$$

The lower and upper bounds of the optimization problem are chosen as physically meaningful ranges for acid pK_a values. Initial values of $pK_{a1,0} = pK_{a2,0} = 7$ are selected. In the case of formic, acetic, and citric acid, the optimization problem, number of DOFs, and underlying equations are adjusted according to the respective acid dissociation equilibria. The optimization problem in Eq. (1) is implemented and solved applying our open-source software for deterministic global optimization MAiNGO (AVT.SVT, RWTH Aachen University).[61–63]

2.3.6 Mixture Hard Model Calibration

For calibration of mixture HMs and PLS models, as well as for spectral data evaluation, PEAX-ACT is used. In the case of itaconic acid (details on other acids in SM), different mixture HMs and PLS models are calibrated employing three different calibration data sets separately: (i) a full set comprising all 22 Raman spectra of the titration experiment with corresponding composition data calculated with estimated pK_a, (ii) a full set comprising the same Raman spectra with composition data that is calculated on the basis of reported pK_a, and (iii) a minimum set

comprising only 4 specific Raman spectra of the titration experiment (maximum molar ratio of IA, IA^-, and IA^{2-}, respectively, and pure water) with their composition data determined from reported pK_a (cf. SM for an alternative selection of a minimum set of calibration samples).[60] The minimum calibration sample number is supported by Alsmeyer et al. and should help to elucidate the effect of sample number on the calibration.[45] In principle, some deviation from the maximum concentration of each acid species is possible, but this reduces the corresponding Raman signal of the acid species and could result in less accurate PCMs as a consequence. Hence, samples with maximum acid species concentration are desired. The mixture HMs are calibrated subject to a closure constraint $\sum x_{k,j} = 1$ with $k \in [\mathrm{IA}, \mathrm{IA}^-, \mathrm{IA}^{2-}, \mathrm{water}]$ that holds for each titration step j. Due to calibration with active closure constraint, the selected maximum function in PEAXACT is of the type simple (ratios) for all HMs. Details on PLS calibration settings are provided in the SM.

2.3.7 Validation of Chemometrics

For all HMs and PLS models, a leave-10%-out cross-validation is done to assess the calibration performance. Figures of merit such as coefficients of determination ${R^2}_k$, root mean square errors of leave-10%-out cross-validation ($RMSECV_k$), and root mean square errors of prediction ($RMSEP_k$) for each chemical species k are calculated to assess the performance of the chemometric models (HM and PLS).[42]

The limit of detection $LOD_k = \bar{x}_{k,\mathrm{blank}} + \beta\sigma_{k,\mathrm{blank}}$ is determined for each chemical species k by the evaluation of 10 blank measurements in pure deionized water. The mean mole fraction $\bar{x}_{k,\mathrm{blank}}$ and corresponding standard deviation $\sigma_{k,\mathrm{blank}}$ are derived from the model prediction of the blank measurements. A confidence factor of $\beta = 3.3$ is chosen that corresponds to a confidence level of 95% assuming normal distribution of measurement errors.[64]

3 Results and Discussion

3.1 Inline Raman Spectra of Itaconic Acid Titration

In Fig. 2, a set of inline Raman spectra recorded during full titration of itaconic acid with NaOH in aqueous solution is depicted. For the assessment of the pK_a estimation versus the known pK_a, a set of 4 calibration samples is taken from the full set consisting of 22 calibration samples. As schematically shown in the zoomed fingerprint region in Fig. 2, itaconic acid dissociates in two stages in course of the titration from fully protonated IA over the first dissociation stage IA^- to the second, fully de-protonated stage IA^{2-}. The corresponding dissociation equilibria are determined by the values for pK_{a1} and pK_{a2}.

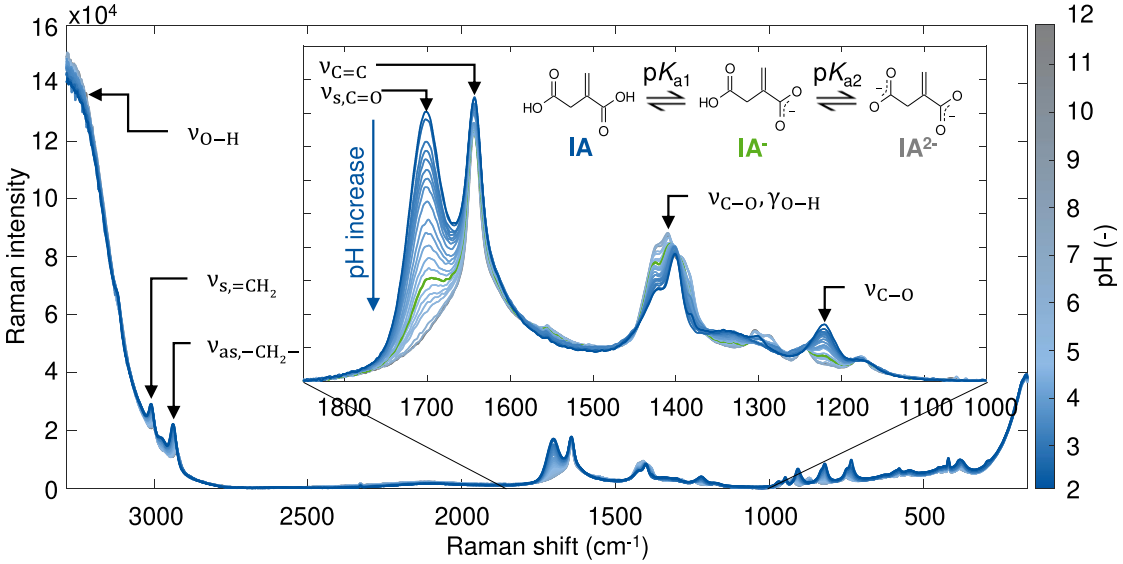

Figure 2. Inline Raman spectra of itaconic acid in aqueous solution, titrated with NaOH at 25 °C by 22 titration steps. The zoom details the fingerprint region.

The strongest Raman bands occur at a high wavenumber from 2800 – 3285 cm^{-1} for $\nu_{O\text{-}H}$ stretching vibrations predominately caused by water. Further, $\nu_{s,=CH_2}$ symmetric stretching modes at 3010 cm^{-1} and $\nu_{s,\text{-}C_2H\text{-}}$ asymmetric stretching modes at 2939 cm^{-1} belong to itaconic acid.[65] The most prominent vibrational bands of itaconic acid visible in the fingerprint region

are the carbonyl symmetric stretching mode $\nu_{s,C=O}$ at 1700 cm^{-1}, the carbon-carbon double bond stretching mode $\nu_{C=C}$ at 1643 cm^{-1}, the combined $\nu_{C\text{-}O}$ stretching and $\gamma_{\text{-O-H}}$ deformation modes around 1400 cm^{-1}, and the $\nu_{C\text{-}O}$ stretching mode of the carboxyl group at 1220 cm^{-1}.[65]

With increasing pH, a progressive decrease of the carbonyl stretching mode $\nu_{s,C=O}$ at 1700 cm^{-1} in conjunction with a decrease of the stretching mode $\nu_{C\text{-}O}$ at 1220 cm^{-1} is observed. The peak decrease is caused by the increasing dissociation of itaconic acid, which leads to a successively decreasing number of carboxylic groups exhibiting the respective vibrational modes. The resulting carboxylate ion peaks are only slightly visible in a spectral range of 1400 – 1600 cm^{-1}.[66]

The $\nu_{C\text{-}O}$ stretching and $\gamma_{\text{-O-H}}$ deformation modes around 1400 cm^{-1} slightly shift towards a higher wavenumber by 8.8 cm^{-1} with increasing pH. The shift can also be explained by the increasing level of itaconic acid dissociation causing a higher C-O bond strength and therefore a higher wavenumber of the molecular vibration. Weak changes in peak intensities and positions occur at a lower wavenumber of 700 – 1000 cm^{-1} but are not evaluated in more detail in this work.

3.2 Partial Least Squares Regression for Itaconic Acid

We establish PLS calibrations based on the titration samples in Fig. 2 and calculated compositions from reported pK_a values as calibration benchmarks because PLS is a widely used standard method for quantitative evaluation of complex spectral data. PLS is calibrated using the same data sets as described in the section Mixture Hard Model Calibration (details on PLS calibration in SM, Tab. SII). For each chemical species k, the minimum calibration set yields varying ${R^2}_k$ and rather high $RMSEP_k$ between 3.16 – 17.06 $\times 10^{-4}$ mol mol^{-1} that are also reflected in the parity plots (Fig. S6, SM), which show strong non-physical trends. The full set performs better but with still relatively high $RMSECV_k$ between 2.13 – 8.41 $\times 10^{-4}$ mol mol^{-1} (cf. Fig. S5, SM). The rather unsatisfying calibration results of PLS can be attributed to the data-driven nature of PLS. Moreover, without a priori available pK_a, calculation of acid dissociation and sample composition fails, therefore rendering the PLS method infeasible at all. To overcome the mentioned

difficulties during PLS calibration, we pursue the suggested methods A and B in Fig. 1 in the following.

3.3 Method A: Monitoring of Acid Species Concentration With Known Dissociation Constants

As outlined in Fig. 1 for the case of known pK_a, the a priori calculation of sample composition with reported pK_{a1} = 3.85 and pK_{a2} = 5.45 enables calculation of the dissociated itaconic acid species IA, IA^-, and IA^{2-} (Fig. 3(a)).[60] The symbols (ˆ) and (˜) indicate spectra with maximum concentration of acid species that are ready to be used for CHM and spectra that need further processing before CHM, respectively. Dissociation of IA proceeds with increasing pH by forming IA^- with a maximum at pH = 4.68, followed by nearly complete dissociation to IA^{2-} at pH = 7 and higher.

Maximum molar ratio of the respective itaconic acid species can be measured at the dashed vertical lines in Fig. 3(a). The first titration step at pH = 2.09 ($\hat{S}_{IA}$) shows a molar ratio of acid species of $x^{rel}_{IA,1}$ = 0.9828 mol mol^{-1} for IA in water. Similarly, only IA^{2-} and water contribute to the last titration step at pH = 11.84 with the corresponding spectrum $\hat{S}_{IA^{2-}}$. While the spectra $\hat{S}_{IA}$ and $\hat{S}_{IA^{2-}}$ can be used right away for PCM construction via CHM, the situation for the spectrum $\tilde{S}_{IA^-}$ at pH = 4.68 is more complicated. All three itaconic acid species contribute to the overall itaconic acid content in water to a significant extent with molar ratios of acid species of $x^{rel}_{IA,15}$ = 0.1133 mol mol^{-1}, $x^{rel}_{IA^-,15}$ = 0.7590 mol mol^{-1}, and $x^{rel}_{IA^{2-},15}$ = 0.1277 mol mol^{-1}, where the subscript number refers to the titration step 15. The coexistence of all three structurally similar itaconic acid species for the titration step 15 at pH = 4.68 impedes the use of the raw spectrum $\tilde{S}_{IA^-}$ for direct application in CHM to construct a PCM of IA^-. To overcome this problem, the spectra $\hat{S}_{IA}$ and $\hat{S}_{IA^{2-}}$ are weighted on the basis of calculated concentration data and subtracted from measured raw spectrum $\tilde{S}_{IA^-}$ (details provided in the SM, Eq. (S16)-(S17)). This results in a processed spectrum $\hat{S}_{IA^-}$, shown in Fig. 3(b), which corresponds to the IA^- Raman signal, still with underlying water signal.

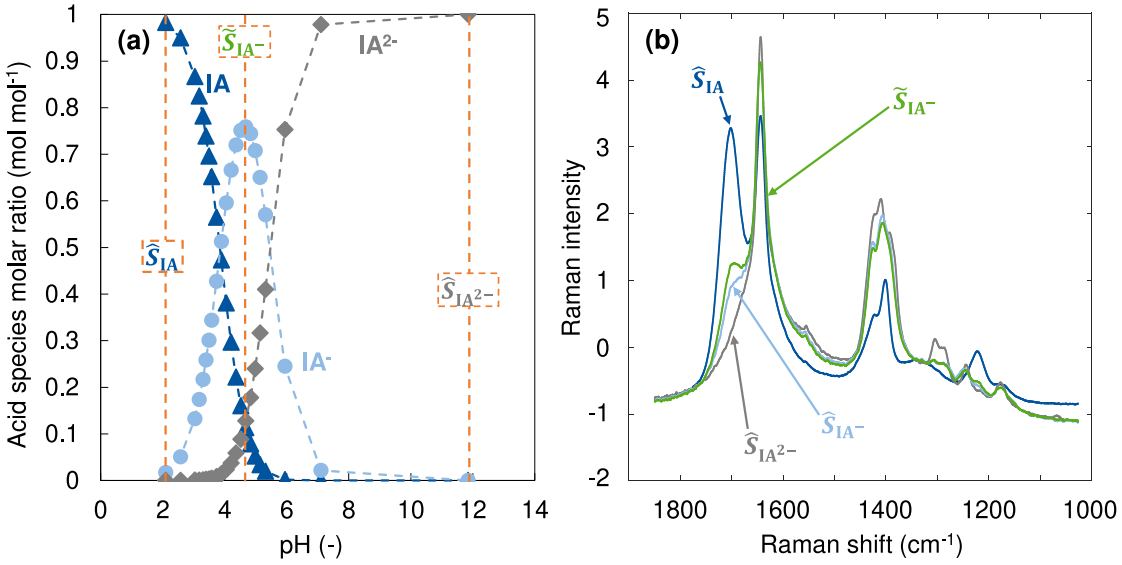

Figure 3. **(a)** Dissociated itaconic acid species calculated from different pH with Raman spectra in Fig. 2. Vertical dashed lines indicate concentrations with maximum molar ratio for each of the three itaconic acid species. **(b)** Raman spectra of maximum molar ratios in **(a)** and processed (ˆ) spectrum $\hat{S}_{\mathrm{IA}^-}$ resulting from weighted linear subtraction of $\hat{S}_{\mathrm{IA}}$ and $\hat{S}_{\mathrm{IA}^{2-}}$ from the raw (˜) spectrum $\tilde{S}_{\mathrm{IA}^-}$.

The most significant difference of $\tilde{S}_{\mathrm{IA}^-}$ and $\hat{S}_{\mathrm{IA}^-}$ emerges at 1700 cm^{-1} where the carbonyl symmetric stretching mode $\nu_{\mathrm{s,C=O}}$ is reduced in intensity, because fully protonated IA does not contribute to this peak anymore. All other parts of the spectrum $\hat{S}_{\mathrm{IA}^-}$ differ only slightly from $\tilde{S}_{\mathrm{IA}^-}$, because IA^- is already the main acid species contributing to the overall itaconic acid Raman signal. This analysis after spectral subtraction is in line with the qualitative observations in Fig. 2. The processed spectrum $\hat{S}_{\mathrm{IA}^-}$ now enables construction of a PCM for IA^- via CHM that is straightforward.

The PCMs of all three itaconic acid species and water that are constructed in the second major task according to Fig. 1 are depicted in Fig. 4(a)–(d). The mixture HM comprising all four PCMs is fitted to a spectrum of an aqueous solution containing all three species of itaconic acid and NaOH at pH = 4.83 and is shown in Fig. 4(e) that represents a typical quality of the HM fit

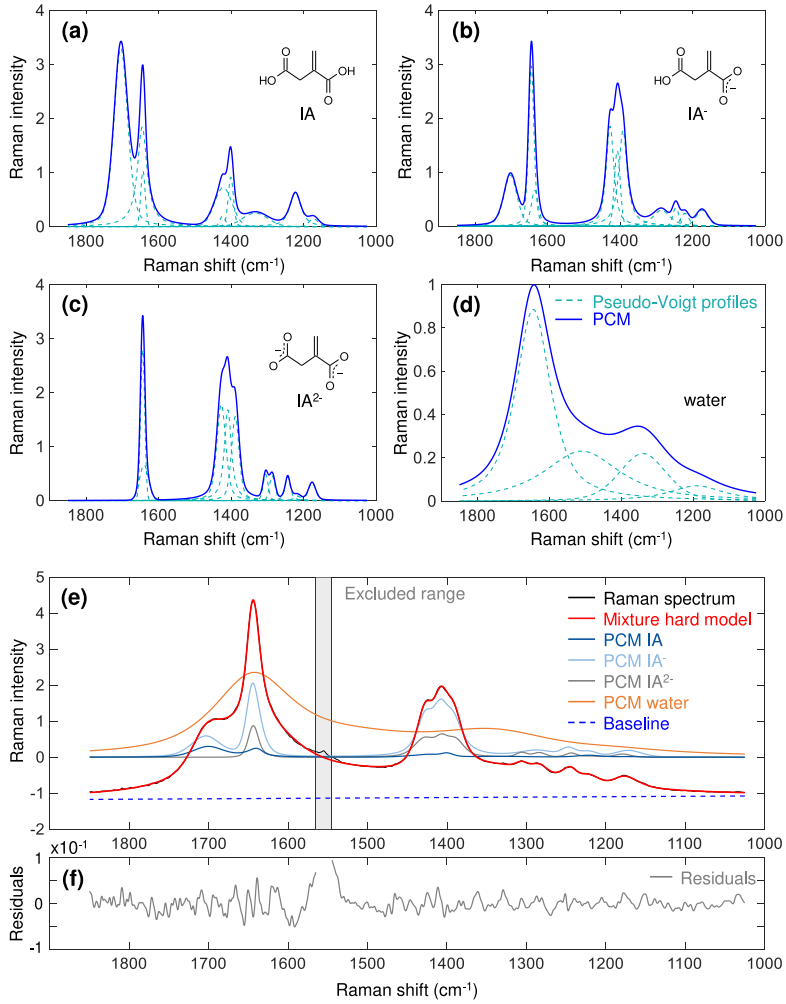

Figure 4. **(a)**–**(d)** PCMs (——) comprised of pseudo-Voigt profiles (- - -) for the itaconic acid species IA, IA^-, IA^{2-}, and water, based on linearly weighted spectra subtraction (known pK_a), without weighting. **(e)** Mixture HM (——) comprised of PCMs from **(a)**–**(d)** with excluded range and linear baseline (- - -) for IA, IA^-, IA^{2-}, and water fitted to a Raman spectrum (——) of an aqueous solution containing itaconic acid and NaOH at pH = 4.83. **(f)** Residuals of spectral fit.

(cf. Fig. S3 in SM for *RMS Res* over pH). The model reflects the spectral data very well in the selected spectral range as indicated by the low and uniformly distributed spectral residuals in Fig. 4(f).

356 In the fourth major task in Fig. 1, the mixture HM is calibrated and validated. Calibration
357 of the mixture HM using the minimum set (4 samples) and the full set (22 samples) yields
358 the figures of merit that are listed in Tab. I. In the case of the minimum set, judgement of the
359 calibration performance by evaluation of ${R^2}_k$ and $RMSECV_k$ fails because these figures of merit
are strongly biased by the low number of calibration samples.

Table I. Figures of merit for HM calibration of itaconic acid dissociation in aqueous solution for each chemical species k visible in the Raman calibration spectra.

Chemical species k	${R^2}_k$ -	$RMSECV_k$	$RMSEP_k$	LOD_k
		$\times 10^{-4}$ mol mol^{-1}		
(IA)[a]	0.9999	-	2.87	1.20
(IA)[b]	0.9974	1.70	-	1.26
(IA^-)[a]	0.9984	-	2.32	0.37
(IA^-)[b]	0.9826	2.13	-	0.37
(IA^{2-})[a]	0.9993	-	2.17	0.39
(IA^{2-})[b]	0.9819	2.02	-	0.37
(water)[a]	0.9989	-	2.50	-
(water)[b]	0.9942	1.70	-	-

[a] 4 calibration samples (min. set) and 18 validation samples.

[b] 22 calibration samples (full set) and cross-validation.

360

361 A more suitable metric to evaluate the model performance for the minimum set is the $RMSEP_k$
362 that we calculate by employing the remaining 18 titration Raman spectra from Fig. 3(a) as a
363 validation set. Calculation of $RMSEP_k$ yields low values of less than 2.87×10^{-4} mol mol^{-1} for

all acid species and water. Values for LOD_k are in the same range or even one order of magnitude lower than the corresponding $RMSEP_k$.

In comparison to the PLS calibration using the minimum set, the HM calibration shows superior performance with a lower $RMSEP_k$ and LOD_k up to one order of magnitude. This also becomes apparent by comparison of the HM and PLS parity plots (cf. Fig. S6, SM) where strong non-physical trends can be observed for the validation sample in the case of data-driven PLS, whereas the physically-motivated HM predicts the validation sample composition accurately even with the minimum set of calibration samples.

For the full set, all figures of merit but $RMSEP_k$ are evaluated: ${R^2}_k$ values are close to 1, indicating that the model can explain the variance of the underlying calibration data very well. The values for $RMSECV_k$ are small and equal to or below 2.13×10^{-4} mol mol^{-1} for all acid species and water. All $RMSECV_k$ values are in the range of or one order of magnitude lower than the minimum species content in the calibration samples (cf. Fig. S4, SM), which enables the envisioned monitoring of acid species. Likewise, the values for LOD_k are in the same range or even one order of magnitude lower than the corresponding $RMSECV_k$. Similar to the minimum set, the HM calibration using the full set yields better results compared to the respective PLS calibration. This proves the physical motivation of the HM, which results in more consistent spectra evaluation.

Comparison of the minimum and the full calibration set shows that $RMSEP_k$ values from the minimum set are only slightly higher than the values for $RMSECV_k$ in the case of the almost six times larger full set. This nicely demonstrates the applicability of IHM on a very small calibration data set. Our study shows that the calibration effort for a HM can be strongly reduced when samples with maximum concentration of each species can be prepared.

3.4 Method B: Monitoring of Acid Species Concentration for Unknown Dissociation Constants

We now hypothetically assume unknown acid dissociation constants of itaconic acid. As a consequence, four major tasks have to be fulfilled for the construction and calibration of a quantitative mixture HM (see Fig. 1). First, two-species spectra of each itaconic acid species in water need to be generated for PCM construction. Second, a mixture HM needs to be constructed from the PCMs. Third, acid dissociation constants in terms of pK_{a1} and pK_{a2} need to be estimated for the quantification of calibration sample composition that enables calibration and validation in the fourth major task.

Because of unknown pK_a, identification of the necessary mixture spectra for PCM construction is impossible on the basis of calibration sample composition and weighted spectra subtraction as in the case of known pK_a (cf. Fig. 1). To still obtain spectra of each itaconic acid species in water, the Raman spectra in Fig. 2 are processed by MCR-ALS subject to settings and constraints outlined in the methods section. Calculation of a water spectrum with MCR-ALS is difficult due to the low Raman activity of water and strong overlap of water with acid vibrational bands. Therefore, we do not use MCR-ALS to calculate the underlying pure component spectra of all four chemical species (three acid species plus water). Instead, we calculate only three spectra of which each spectrum comprises signals of only one acid species together with signals of water to some degree. Hence, we calculate two-species spectra by MCR-ALS. As the corresponding component concentration data from MCR-ALS (simultaneously obtained with the two-species spectra for each titration sample) is also impaired by a latent water content, the direct use of this data for pK_a estimation is prohibited and only possible with IHM.

The resulting spectra are shown in Fig. 5 for the three itaconic acid species IA, IA^-, and IA^{2-}. Despite no input about spectral characteristics, the MCR-ALS results are very similar to the two-species spectra obtained from weighted linear spectra subtraction in Fig. 3. The prevalent difference is a two-species spectrum for IA^- in water by MCR-ALS with higher Raman intensity of the carbonyl stretching mode $\nu_{s,C=O}$ at 1700 cm^{-1} relative to the carbon-carbon double bond

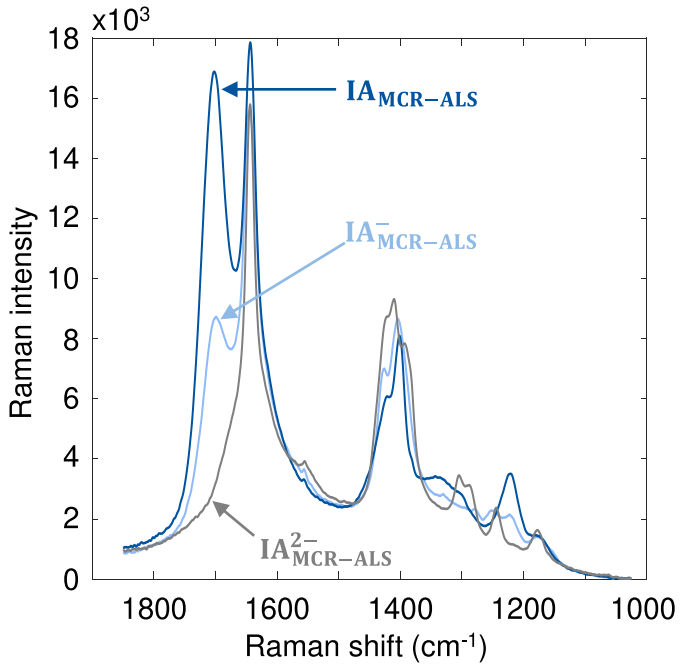

Figure 5. Two-component spectra of itaconic acid species IA, IA^-, and IA^{2-} in water, obtained by MCR-ALS evaluation of Raman spectra in Figure 2, for PCM construction during IHM (cf. Fig. S10, SM, for PCM comparison of method A and B).

stretching mode $\nu_{C=C}$ at 1643 cm^{-1} as compared to the spectrum in Fig. 3. The reason for this could be the unknown extent of the water signal in the resulting two-species spectrum, which increases the intensity in the corresponding wavenumber interval. As CHM is used for PCM construction (direct comparison of PCMs resulting from method A and B in Fig. S10, SM), two-species spectra can be applied right away and IHM in the second main task is straightforward (cf. Fig. S8, SM, for PCMs, mixture HM and fitting accuracy).

To assess the third major task of composition calculation via estimating pK_{a1} and pK_{a2}, the dissociation equilibria of itaconic acid are fitted to the ratios of the peak areas of itaconic acid species from the HM that is applied on the titration spectra in Fig. 2, with pK_{a1} and pK_{a2} as DOF. The fitting results are presented in Fig. 6. In general, the estimated model nearly perfectly fits the true values. Observed deviations originate from inconsistent spectral fits of the mixture HM

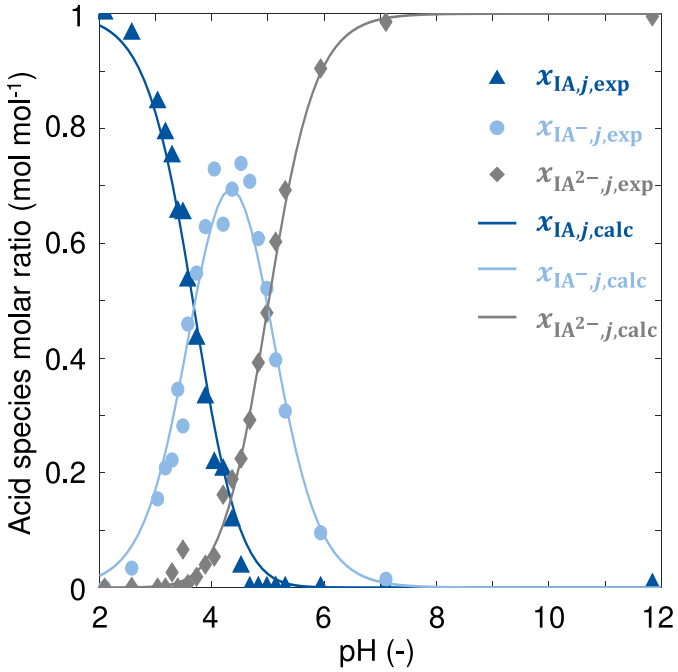

Figure 6. Itaconic acid dissociation equilibria and species balance fitted to the ratios of the peak areas of IA species from IHM to determine pK_{a1} and pK_{a2}. Estimation of itaconic acid dissociation constants yields pK_{a1} = 3.68 and pK_{a2} = 5.00.

to the titration Raman spectra (probably due to measurement noise and difficulty of fitting the weak water Raman signal).

The determination of acid dissociation constants is commonly done at constant ionic strength I and temperature T to avoid changing activity coefficients. In the present work, T is kept constant at 25 °C, but I changes with a mean value of $I = 0.34$ mol L^{-1} and influences are neglected. The estimation of pK_{a1} and pK_{a2}, based on the curve fitting shown in Fig. 6, yields values of pK_{a1} = 3.68 and pK_{a2} = 5.00. The values of pK_{a1} and pK_{a2} are consistent with the range of acid dissociation constants for itaconic acid found in the literature ($3.48 \leqslant pK_{a1} \leqslant 3.95$ and $4.98 \leqslant pK_{a2} \leqslant 5.67$, cf. Tab. SIV, SM). With the estimated pK_a values, the mole fractions of IA, IA$^-$, IA^{2-}, and water can be determined for the titration steps corresponding to the 22 Raman spectra in Fig. 2 to complete the calibration data set. It should be noted that until here,

no calibration of the mixture HM is done, but values for pK_a are identified on a calibration-free basis. This makes method B independent of a priori known dissociation constants and holds potential for other applications that are discussed at the end of this section.

In the fourth major task, the mixture HM is calibrated using the full set employing the estimated pK_a values. The results are summarized in Tab. II (parity plots provided in Fig. S11, SM).

Table II. Figures of merit for HM calibration of itaconic acid dissociation in aqueous solution for each chemical species k in the full set of 22 calibration samples. Estimated values of pK_{a1} = 3.68 and pK_{a2} = 5.00 are employed for calculation of the sample composition.

Chemical species k	${R^2}_k$ -	$RMSECV_k$ $\times 10^{-4}$ mol mol^{-1}	LOD_k
IA	0.9927	2.40	1.29
IA^-	0.9793	2.26	0.45
IA^{2-}	0.9946	1.12	0.46
water	0.9849	2.26	-

All ${R^2}_k$ values are close to 1 indicating that the model fits the variance of the underlying calibration data well. The slightly lower value of ${R^2}_{IA^-}$ corresponds to the less accurate fit of the mixture HM to the Raman spectra in the range of pH = 3 to pH = 5 visible in Fig. 6. The low values for $RMSECV_k$ confirm the accuracy of the calibration. Values for $RMSECV_k$ are only slightly higher compared to the full set calibration results in Tab. I, which supports the consistency of the two approaches.

Values for LOD_k are in the same range or even one order of magnitude lower than the corresponding $RMSECV_k$ and indicate a sufficiently precise and consistent calibration model.

The herein presented method B of coupling IHM and MCR-ALS is independent of a priori

available thermodynamic data such as dissociation constants in the present case. This aspect holds further potential as the approach can be transferred to other problems such as estimation of kinetic parameters from spectroscopic data. By coupling the HM fitting of the spectral data to the estimation of thermodynamic parameters, a coupled optimization problem is formed that covers the mass balance not just in one spectrum (physically justified HMs) but with the pK_a over the whole spectral data set. The optimization task is simultaneous minimization of the spectral residuals and estimation of the pK_a.

3.5 Generalizability of the Methods

To demonstrate the generalizability of our method A to other carboxylic acids, mixture HMs are constructed and calibrated for formic, acetic, and citric acid for known pK_a.[69–71] Furthermore, we apply method B to formic and acetic acid to validate its applicability to other carboxylic acids. In the herein presented case studies, the assumption of a dilute ideal mixture does not hold strictly speaking, because acid concentrations surpass 0.1 mol L^{-1}, ionic strength exceeds 0.01 mol L^{-1} and polyvalent ions occur. This requires to consider activities for thermodynamically correct calculations.[67,68] Nevertheless, our experience shows that our strategy can be applied with good accuracy to different carboxylic acids. For these reasons, activities are not explicitly considered in the present work and ideal thermodynamics are used instead.

The calibration results for formic, acetic, and citric acid according to method A are summarized in Tab. III (cf. SM for modeling details and parity plots).

Table III. Figures of merit for HM calibration of formic, acetic, and citric acid dissociation in aqueous solution for each chemical species k visible in the Raman calibration spectra according to method A. Reported pK_a values are employed for calculation of calibration sample composition.[69–71]

Chemical species k	${R^2}_k$ -	$RMSECV_k$ $\times 10^{-4}$ mol mol^{-1}	LOD_k
FA	0.9982	2.61	0.35
FA^-	0.9946	2.29	0.43
water	0.9979	1.93	-
AA	0.9960	2.43	0.00
AA^-	0.9966	2.09	0.00
water	0.9946	1.79	-
CA	0.9981	2.99	0.00
CA^-	0.9675	7.16	12.99
CA^{2-}	0.9225	8.81	7.51
CA^{3-}	0.9154	7.20	2.96
water	0.9463	10.13	-

The calibrations of formic and acetic acid via method A yield accurate results proven by low values for $RMSECV_k$ that are close to the results of the itaconic acid calibration. In the case of citric acid, some difficulties arise during calibration using method A. As the molecular structure and therefore PCMs of CA^{2-} and CA^{3-} are very similar, they can hardly be distinguished by the IHM algorithm. This leads to lower HM calibration accuracies and higher validation errors.

The results of method B applied on formic and acetic acid are provided in the SM, Fig. S15 and Fig. S19. We estimate values of $pK_a = 3.59$ for formic and $pK_a = 4.70$ for acetic acid that are both close to reported values.[69,70] The figures of merit for HM calibration are very comparable to the results of method A for both acids and are even slightly better (cf. SM, Fig. S15 and Fig. S19).

Moreover, the comparison of the HM and PLS calibrations for formic, acetic, and citric acid reveals errors that are up to one order of magnitude larger in the case of PLS with $RMSECV_k$ ranging from 5.94 – 23.04 mol mol^{-1} as shown in Fig. IV (details on calibration settings in Tab. SVI, SM). The general performance of PLS could be improved by using more sophisticated PLS approaches that are pursued in future work.[43,44]

Table IV. Figures of merit for PLS calibration of formic, acetic, and citric acid dissociation in aqueous solution for each chemical species k visible in the Raman calibration spectra. Reported pK_a values are employed for calculation of calibration sample composition.[69–71]

Chemical species k	${R^2}_k$ -	$RMSECV_k$ $\times 10^{-4}$ mol mol^{-1}	LOD_k
FA	0.9981	12.01	2.82
FA^-	0.9900	5.94	0.21
water	0.9998	6.23	-
AA	0.9993	10.67	22.12
AA^-	0.9970	11.06	16.56
water	0.9992	21.78	-
CA	0.9971	8.41	14.83
CA^-	0.9697	17.81	12.63
CA^{2-}	0.8665	23.04	51.66
CA^{3-}	0.8250	16.34	36.43
water	0.9983	19.30	-

We conclude that the methods A and B are generally applicable for quantitative concentration monitoring of acid dissociation states.

The combination of inline Raman spectroscopy and IHM selectively reveals the dissociation species concentration profiles with good accuracy at low experimental calibration effort and therefore allows for quantitative process monitoring. Our approach is theoretically not limited to carboxylic acids but can be applied to Raman-active mineral acids and acid mixtures also at higher concentrations. In such cases, attention has to be paid to the relevant assumptions. Moreover, Raman-active electrolytes can be calibrated and quantified by our approach that, hence, bears potential to support analytics for real-time observation of electrochemical reactions and processes.[72,73]

The presented approach of Raman spectroscopy, IHM, and model-based calibration enables reliable monitoring of the concentrations of acid species at elevated temperatures, in solutions of higher ionic strength or multiple pH-active chemicals, where in-depth information on acid dissociation is not accessible by conventional methods such as pH measurements.

4 Conclusion

We present a physically justified approach for the quantification of dissociated carboxylic acid species in aqueous solution on the basis of Raman spectroscopy and IHM for the cases of both known (method A) and unknown (method B) acid dissociation constants pK_a. A comparison of IHM with data-driven PLS results for calibration proves much worse errors or even non-reliable PLS calibrations. Calibration of itaconic acid species in water via method A by IHM yields very good results with $RMSECV_k$ values between 1.70 – 2.13×10^{-4} mol mol^{-1} for 22 calibration samples and slightly worse $RMSEP_k$ values of less than 2.87×10^{-4} mol mol^{-1} even with 4 calibration samples only. This enables a very efficient calibration strategy as only a small set of calibration samples is required. In the case of unknown acid dissociation constants, method B can also estimate the pK_a of the acid equilibrium, which enables accurate and consistent process

monitoring of acids in aqueous solutions. With itaconic acid, the values of pK_{a1} = 3.68 and pK_{a2} = 5.00 are estimated (compared to pK_{a1} = 3.85 and pK_{a2} = 5.45 from the literature). HM errors of $RMSECV_k$ between 1.12 – 2.40$\times 10^{-4}$ mol mol^{-1} are reached. Successful application of both methods to formic and acetic acid and method A to citric acid in aqueous solution with similar $RMSECV_k$ between 1.14 – 10.13$\times 10^{-4}$ mol mol^{-1} verifies the generalizability of the suggested methods for other Raman-active acids.

5 Acknowledgements

The authors thank Felix Berger, Nina Stein, Niklas Rische and Hannah Ingendae for their valuable support.

6 Declaration of Conflicting Interests

The authors declare that there is no conflict of interest.

7 Funding

The authors gratefully acknowledge the financial support of the Project ContiHighSolid by the Federal Ministry of Education and Research (BMBF) and the project supervision by the project management organization Projektträger Jülich (PtJ) that enabled the presented research. Furthermore, parts of this work were performed within the Cluster of Excellence 236 “Tailor-Made Fuels from Biomass”, which is funded by the Excellence Initiative by the German federal and state governments to promote science and research at German universities. Additionally, parts of this work were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy - Exzellenzcluster 2186 “The Fuel Science Center”.

8 Supplemental Material

All supplemental material mentioned in the text, including experimental setup, equations for dissociation equilibrium calculation, PCMs, mixture HMs, parity plots and figures of merit, is available in the online version of the journal.

9 ORCID iD

Alexander Echtermeyer: https://orcid.org/0000-0002-9382-2227
Caroline Marks: https://orcid.org/0000-0002-0917-7051
Alexander Mitsos: https://orcid.org/0000-0003-0335-6566
Jörn Viell: https://orcid.org/0000-0003-0587-6151

10 Author Contributions

- **(AE)**: Acquisition of inline Raman data for formic acid and citric acid titration; IHM calibration and application for quantitative analysis; MCR-ALS analysis; Formulation of the parameter estimation problem; Estimation of acid dissociation constants; Evaluation of data; Discussion on method and results; Preparation of manuscript
- **(CM)**: Acquisition of inline Raman data for acetic acid titration; Analysis of aspects influencing the proposed analytical method; Discussion on method and results; Preparation of manuscript
- **(AM)**: Design of the project; Scientific support; Formulation of the parameter estimation problem; Advice on structure and presentation of this work; Reviewing and editing of the manuscript
- **(JV)**: Design of the project; Acquisition of inline Raman data for itaconic acid titration; Scientific support, guidance, and discussion on method and results; Advice on structure

and presentation of this work; Reviewing and editing of the manuscript

References

[1] A.M. Johnson, H. Kim, J. Ralph, S.D. Mansfield. “Natural Acetylation Impacts Carbohydrate Recovery During Deconstruction of *Populus trichocarpa* Wood". Biotechnol. Biofuels. 2017. 10(48): 1-12.

[2] D. Di Marino, T. Jestel, C. Marks, J. Viell, M. Blindert, S.M.A. Kriescher, A.C. Spiess, M. Wessling. “Carboxylic Acids Production via Electrochemical Depolymerization of Lignin". ChemElectroChem. 2019. 6(5): 1434-1442.

[3] L.T. Mika, E. Cséfalvay, A. Németh. “Catalytic Conversion of Carbohydrates to Initial Platform Chemicals: Chemistry and Sustainability". Chem. Rev. 2018. 118(2): 505-613.

[4] S.G. Maerten, D. Voß, M.A. Liauw, J. Albert. “Selective Catalytic Oxidation of Humins to Low-Chain Carboxylic Acids With Tailor-Made Polyoxometalate Catalysts". ChemistrySelect. 2017. 2(24): 7296-7302.

[5] F. Valentini, V. Kozell, C. Petrucci, A. Marrocchi, Y. Gu, D. Gelman, L. Vaccaro. “Formic Acid, a Biomass-Derived Source of Energy and Hydrogen for Biomass Upgrading". Energy Environ. Sci. 2019. 12(9): 2646-2664.

[6] M. Aigner, D. Roth, J. Rußkamp, J. Klankermayer, A. Jupke. “Model-Based Equipment Design for the Biphasic Production of 5-Hydroxymethylfurfural in a Tubular Reactor". AIChE J. 2019. 66(4): e16849.

[7] L. Karaffa, C.P. Kubicek. “Citric Acid and Itaconic Acid Accumulation: Variations of the Same Story?". Appl. Microbiol. Biotechnol. 2019. 103(7): 2889-2902.

[8] T. Klement, J. Büchs. “Itaconic Acid - A Biotechnological Process in Change". Bioresour. Technol. 2013. 135: 422-431.

[9] J.J. Bozell, G.R. Petersen. “Technology Development for the Production of Biobased Products From Biorefinery Carbohydrates – the US Department of Energy’s “Top 10” revisited". Green Chem. 2010. 12(4): 539-554.

[10] T. Werpy, G. Petersen. “Top Value Added Chemicals From Biomass: Volume I – Results of Screening for Potential Candidates From Sugars and Synthesis gas". Golden, USA: National Renewable Energy Lab, 2004.

[11] R. Bafana, R.A. Pandey. “New Approaches for Itaconic Acid Production: Bottlenecks and Possible Remedies". Crit. Rev. Biotechnol. 2018. 36(1): 68-82.

[12] A.I. Magalhães, J.C. De Carvalho, J.D.C. Medina, C.R. Soccol. “Downstream Process Development in Biotechnological Itaconic Acid Manufacturing". Appl. Microbiol. Biotechnol. 2017. 101(1): 1-12.

[13] J. Gorden, E. Geiser, N. Wierckx, L.M. Blank, T. Zeiner, C. Brandenbusch. “Integrated Process Development of a Reactive Extraction Concept for Itaconic Acid and Application to a Real Fermentation Broth". Eng. Life Sci. 2017. 17(7): 809-816.

[14] A. Eggert, T. Maßmann, D. Kreyenschulte, M. Becker, B. Heyman, J. Büchs, A. Jupke. “Integrated In-Situ Product Removal Process Concept for Itaconic Acid by Reactive Extraction, pH-Shift Back Extraction and Purification by pH-Shift Crystallization". Sep. Purif. Technol. 2019. 215: 463-472.

[15] C. Kocks, J. Görtz, A. Holtz, M. Gausmann, A. Jupke. “Electrochemical Crystallization Concept for Succinic Acid Reduces Waste Salt Production". Chem. Ing. Tech. 2020. 92(3): 221-228.

[16] M. Gausmann, C. Kocks, M. Doeker, A. Eggert, T. Maßmann, A. Jupke. “Recovery of Succinic Acid by Integrated Multi-Phase Electrochemical pH-Shift Extraction and Crystallization". Sep. Purif. Technol. 2020. 240: 116489.

600 [17] K. Eisen, T. Eifert, C. Herwig, M. Maiwald. “Current and Future Requirements to Industrial
601 Analytical Infrastructure - Part 1: Process Analytical Laboratories". Anal. Bioanal. Chem.
602 2020. 412(9): 2027-2035.

603 [18] C. Minnich, S. Hardy, S. Krämer. “Stopping the Babylonian Confusion: An Updated Nomen-
604 clature for Process Analyzers in PAT Applications". Chem. Ing. Tech. 2016. 88(6): 694-697.

605 [19] A.G. Demesa, A. Laari, E. Tirronen, I. Turunen. “Comparison of Solvents for the Recovery
606 of Low-Molecular Carboxylic Acids and Furfural From Aqueous Solutions". Chem. Eng.
607 Res. Des. 2015. 93: 531-540.

608 [20] B. Schrader, D. Bougeard. Infrared and Raman Spectroscopy: Methods and Applications.
609 Weinheim, Germany: Wiley-VCH, 1995. 1st ed.

610 [21] R.W. Kessler. Prozessanalytik: Strategien und Fallbeispiele aus der Industriellen Praxis.
611 Weinheim, Germany: Wiley-VCH, 1995. 1st ed.

612 [22] D. Fraenkel. “Structure and Ionization of Sulfuric Acid in Water". New J. Chem. 2015. 39(7):
613 5124-5136.

614 [23] W.W. Rudolph. “Raman-Spectroscopic Measurements of the First Dissociation Constant of
615 Aqueous Phosphoric Acid Solution From 5 to 301 °C". J. Solution Chem. 2012. 41(4): 630-
616 645.

617 [24] W.W. Rudolph. “Raman- and Infrared-Spectroscopic Investigations of Dilute Aqueous Phos-
618 phoric Acid Solutions". Dalton Trans. 2010. 39(40): 9642-9653.

619 [25] D.A. Knopf, B.P. Luo, U.K. Krieger, T. Koop. “Thermodynamic Dissociation Constant of the
620 Bisulfate ion From Raman and ion Interaction Modeling Studies of Aqueous Sulfuric Acid
621 at low Temperatures". J. Phys. Chem. A. 2003. 107(21): 4322-4332.

622 [26] J.J. Max, C. Ménichelli, C. Chapados. “Infrared Titration of Aqueous Sulfuric Acid". J. Phys.
623 Chem. A. 2000. 104(12): 2845-2858.

[27] J.J. Max, C. Chapados. “Infrared Spectroscopy of Aqueous Carboxylic Acids: Malic Acid". J. Phys. Chem. A. 2002. 106(27): 6452-6461.

[28] J.J. Max, C. Chapados. “Infrared Spectroscopy of Aqueous Carboxylic Acids: Comparison Between Different Acids and Their Salts". J. Phys. Chem. A. 2004. 108(16): 3324-3337.

[29] J.J. Max, M. Trudel, C. Chapados. “Infrared Titration of Aqueous Glycine". Appl. Spectrosc. 1998. 52(2): 226-233.

[30] S. Xi, X. Zhang, Z. Luan, Z. Du, L. Li, B. Wang, L. Cao, C. Lian, J. Yan. “A Direct Quantitative Raman Method for the Measurement of Dissolved Bisulfate in Acid-Sulfate Fluids". Appl. Spectrosc. 2018. 72(8): 1234-1243.

[31] T. Langner, A. Rietig, J. Acker. “Raman Spectroscopic Determination of the Degree of Dissociation of Nitric Acid in Binary and Ternary Mixtures With HF and H_2SiF_6". J. Raman Spectrosc. 2019. 51(2): 366-372.

[32] H. Alatalo, J. Kohonen, H. Qu, H. Hatakka, S.P. Reinikainen, M. Louhi-Kultanen, J. Kallas. “In-Line Monitoring of Reactive Crystallization Process Based on ATR-FTIR and Raman Spectroscopy". J. Chemom. 2008. 22(11-12): 644-652.

[33] H. Qu, H. Alatalo, H. Hatakka, J. Kohonen, M. Louhi-Kultanen, S.P. Reinikainen, J. Kallas. “Raman and ATR FTIR Spectroscopy in Reactive Crystallization: Simultaneous Monitoring of Solute Concentration and Polymorphic State of the Crystals". J. Cryst. Growth. 2009. 311(13): 3466-3475.

[34] M.A. Elbagerma, G. Azimi, H.G.M. Edwards, A.I. Alajtal, I.J. Scowen. “In Situ Monitoring of pH Titration by Raman Spectroscopy". Spectrochim. Acta, Part A. 2010. 75(5): 1403-1410.

[35] M.A. Elbagermi, A.I. Alajtal, H.G.M. Edwards, G.H. Azimi, K.D. Verma, I.J. Scowen. “Raman Spectroscopic and Potentiometric Studies of Acidity Level and Dissociation of Citric Acid in Aqueous Solution". J. Appl. Chem. Sci. Int. 2015. 2(1): 1-11.

649 [36] J. Huguenin, S. Ould Saad Hamady, P. Bourson. “Monitoring Deprotonation of Gallic Acid
650 by Raman Spectroscopy". J. Raman Spectrosc. 2015. 46(11): 1062-1066.

651 [37] A. Beer. “Bestimmung der Absorption des Rothen Lichts in Farbigen Flüssigkeiten". Ann.
652 Physik. 1852. 162: 78-88.

653 [38] S. Wold, M. Sjöström, L. Eriksson. “PLS-Regression: a Basic Tool of Chemometrics".
654 Chemom. Intell. Lab. Syst. 2001. 58(2): 109-130.

655 [39] W. Kessler. Multivariate Datenanalyse: Für die Pharma, Bio- und Prozessanalytik. Weinheim,
656 Germany: Wiley-VCH, 2007. 1st ed.

657 [40] K.S. Booksh. “Chemometric Methods in Process Analysis". In: R.A. Meyers, editor. En-
658 cyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation. USA: John
659 Wiley & Sons, 2006. 1st ed.

660 [41] N. Kumar, A. Bansal, G.S. Sarma, R.K. Rawal. “Chemometrics Tools Used in Analytical
661 Chemistry: An Overview". Talanta. 2014. 123: 186-199.

662 [42] J.P. Conzen. Multivariate Kalibration - Ein Praktischer Leitfaden zur Methodenentwicklung
663 in der Quantitativen Analytik. Ettlingen, Germany: Bruker Optik GmbH, 2005. 4th ed.

664 [43] R. Rosipal. “Nonlinear Partial Least Squares: An Overview". In: H. Lodhi, Y. Yamanishi,
665 editors. Chemometrics and Advanced Machine Learning Perspectives: Complex Computa-
666 tional Methods and Collaborative Techniques, New York, USA: Medical Information Science
667 Reference, 2010. Chap. 9, Pp. 169-189.

668 [44] W.S. Cleveland. “Robust Locally Weighted Regression and Smoothing Scatterplots". J. Am.
669 Stat. Assoc. 1997. 74(368): 829-836.

670 [45] F. Alsmeyer, H.J. Koß, W. Marquardt. “Indirect Spectral Hard Modeling for the Analysis of
671 Reactive and Interacting Mixtures". Appl. Spectrosc. 2004. 58(8): 975-985.

[46] E. Kriesten, F. Alsmeyer, A. Bardow, W. Marquardt. “Fully Automated Indirect Hard Modeling of Mixture Spectra". Chemom. Intell. Lab. Syst. 2008. 91(2): 181–193.

[47] E. Kriesten, D. Mayer, F. Alsmeyer, C.B. Minnich, L. Greiner, W. Marquardt. “Identification of Unknown Pure Component Spectra by Indirect Hard Modeling". Chemom. Intell. Lab. Syst. 2008. 93(2): 108-119.

[48] R. Tauler, B. Kowalski, S. Fleming. “Multivariate Curve Resolution Applied to Spectral Data From Multiple Runs of an Industrial Process". Anal. Chem. 1993. 65(15): 2040-2047.

[49] M. Garrido, F.X. Rius, M.S. Larrechi. “Multivariate Curve Resolution - Alternating Least Squares (MCR-ALS) Applied to Spectroscopic Data From Monitoring Chemical Reactions Processes". Anal. Bioanal. Chem. 2008. 390(8): 2059-2066.

[50] A. De Juan, J. Jaumot, R. Tauler. “Multivariate Curve Resolution (MCR). Solving the Mixture Analysis Problem". Anal. Methods. 2014. 6(14): 4964-4976.

[51] J. Felten, H. Hall, J. Jaumot, R. Tauler, A. De Juan, A. Gorzsás. “Vibrational Spectroscopic Image Analysis of Biological Material Using Multivariate Curve Resolution - Alternating Least Squares (MCR-ALS)". Nat. Protoc. 2015. 10(2): 217-240?.

[52] G.A. De la Rosa, L.C. Zheng, J. Vedad, R.Z.B. Desamero. “Structural, Vibrational, and pK_a Determination of Carboxylic Acids Using DFT Calculations and Raman spectroscopy: An Instrumental Analysis Laboratory". In: M.D. Sonntag, editor. Raman Spectroscopy in the Undergraduate Curriculum. Washington, USA: American Chemical Society, 2018. Chap. 6, Pp. 93-114.

[53] Kaiser Optical Systems Inc. “Technical Note 1250: Immersion Optic for Reaction Monitoring". Ann Arbor, USA: Kaiser Optical Systems, 2006.

[54] MATLAB. Version 9.2.0 (R2017a). Natick, USA: The MathWorks Inc., 2017.

[55] PEAXACT. version 4.5. Aachen, Germany: S-PACT GmbH, 2018.

[56] PEAXACT. version 4.7. Aachen, Germany: S-PACT GmbH, 2019.

[57] K. De Gussem, J. De Gelder, P. Vandenabeele, L. Moens. “The Biodata Toolbox for MATLAB". Chemom. Intell. Lab. Syst. 2009. 95(1): 49-52.

[58] S. Kucheryavskiy. “MATLAB Toolbox for Multivariate Data Analysis". Esbjerg, Demark: Department of Chemistry and Bioscience, Aalborg University, 2016.

[59] W.R. Fenner, H.A. Hyatt, J.M. Kellam, S.P.S Porto. “Raman Cross Section of Some Simple Gases". J. Opt. Soc. Am. 1973. 63(1): 73-77.

[60] H.W. Ashton, J.R. Partington. “The Electrometric Titrations of Some Unsaturated Dicarboxylic Acids". Trans. Faraday Soc. 1934. 30: 598-614.

[61] D. Bongartz, J. Najman, S. Sass, A. Mitsos. “MAiNGO - McCormick-Based Algorithm for Mixed-Integer Nonlinear Global Optimization". 2019. https://git.rwth-aachen.de/avt.svt/public/maingo [accessed Jan 20th 2020].

[62] A. Mitsos, B. Chachuat, P.I. Barton. “McCormick-Based Relaxations of Algorithms". SIAM J. Optim. 2009. 20(2): 573-601.

[63] A. Tsoukalas, A. Mitsos. “Multivariate McCormick Relaxations". J. Global Optim. 2014. 59(2-3): 633-662.

[64] E. Desimoni, B. Brunetti. “About Estimating the Limit of Detection by the Signal to Noise Approach". Pharm. Acta Helv. 2015. 6(3): 355-358.

[65] C.Q. Yang, X. Gu. “Polymerization of Maleic Acid and Itaconic Acid Studied by FT-Raman Spectroscopy". J. Appl. Polym. Sci. 2001. 81(1): 223-228.

[66] J. Oomens, J.D. Steill. “Free Carboxylate Stretching Modes". J. Phys. Chem. A. 2008. 112(15): 3281-3283.

[67] A. Albert, E.P. Serjeant. The Determination of Ionization Constants: A Laboratory Manual. New York, Germany: Chapman and Hall, 2012. 3rd ed.

[68] B.G. Cox. Acids and Bases: Solvent Effects on Acid-Base Strength. Oxford, UK: Oxford University Press, 2013. 1st ed.

[69] J. Mollin, Z. Pavelek, J. Navrátilová, A. Recmanová. “Effect of Medium on Dissociation of Carboxylic Acids". Collect. Czech. Chem. Commun. 1985. 50(12): 2670-2678.

[70] P.K. Dasgupta, O. Nara. “Measurement of Acid Dissociation Constants of Weak Acids by Cation Exchange and Conductometry". Anal. Chem. 1990. 62(11): 1117-1122.

[71] D. Barrón, S. Butí, M. Ruiz, J. Barbosa. “Preferential Solvation in the THF–Water Mixtures. Dissociation Constants of Acid Components of pH Reference Materials". Phys. Chem. Chem. Phys. 1999. 1(2): 295-298.

[72] A.K. Schuppert, A.A. Topalov, I. Katsounaros, S.O. Klemm, K. J.J. Mayrhofer. “A Scanning Flow Cell System for Fully Automated Screening of Electrocatalyst Materials". J. Electrochem. Soc. 2012. 159(11): 670-675.

[73] P. Khanipour, M. Löffler, A.M. Reichert, F.T. Haase, K.J.J. Mayrhofer, I. Katsounaros. “Electrochemical Real-Time Mass Spectrometry (EC-RTMS): Monitoring Electrochemical Reaction Products in Real Time". Angew. Chem. 2019. 131(22): 7351-7355.