Bayesian separation algorithm of THz spectral sources applied to D-glucose monohydrate dehydration kinetics

Lukasz A. STERCZEWSKI[*]
`lukasz.sterczewski@pwr.edu.pl`
Michal P. GRZELCZAK, Kacper NOWAK, Boguslaw SZLACHETKO, Rafal ZDUNEK, and Edward F. PLINSKI
Wroclaw University of Technology - Wybrzeze Wyspianskiego 27 – 50-370 – WROCŁAW - POLAND

Summary: We report the use of hierarchical Bayesian separation algorithm to estimate the kinetics of a polycrystalline hydrate dehydration process. Instead of estimation of the component concentration based on a peak area, a full spectral information in the THz band is taken into account. The components are unmixed by a blind-source separation algorithm. Derived mixing coefficients evolution profile can be used to obtain the dehydration rate of monohydrated D-glucose.

1 Introduction

Hydration is an important class of pseudo-polymorphic transitions of a solid-state [1]. It influences many properties of the substance [2]. THz spectroscopy can support the analysis of the hydration rate, however the estimation of the hydrate and anhydrate forms concentrations may be complex in a presence of other compounds. We investigated a well-known polycrystalline hydrate – D-glucose with evident spectral lines in the THz band. Conventionally, a peak-area or peak height method is used in quantitative analysis. In opposite, we propose an application of the Bayesian BSS algorithm, which provides an information of compounds concentration changing in time, caused by temperature.

2 Method

The D-(+)-glucose monohydrate was diluted with polyethylene powder at 90%, and compressed into a 400 mg pellet (40/360 mg) with a hydraulic press (2 tons pressure for 2 minutes). Its thickness – appr. 3 mm was chosen to prevent a fringe pattern created by an etalon and provide low scattering and the resulting sloping baseline [3]. The pellets (13 mm in diameter) were placed in a custom built Teflon-insulated pharmaceutical samples heater. The temperature remained constant with an accuracy of ±0.5°C. Each D-(+)-glucose monohydrate was heated for 80 minutes at 40°C, 42°C, 44°C, 46°C and 48°C respectively, being measured in a THz-TDS technique concurrently. A femtosecond pulse (780 nm, 85 fs, 15 mW) was focused on a receiving and transmitting photoconductive antenna (PCA). The transmitting antenna was biased with a 20 kHz 30 V_{pp} square wave. The THz pulse was detected in a homodyne coherent detection technique. The sample was placed in the focal point of the second and third parabolic mirror, guiding the THz wave. The experiments were conducted in a pure nitrogen atmosphere, when no water vapour absorption lines were present. A pure PE pellet pulse was considered the reference measurement.

The THz waveforms were obtained by nearly a 3-minute signal accumulation. Fast Fourier Transform was applied to the zero-padded signals. Absorbance was calculated according to the PE reference. Its evolution with time caused by the dehydration process and a time-of-measurement information can be used to determine the dehydration rate. One of such method is a peak-area method, which estimates the concentration of selected spectral source on an area under curve in a selected spectral region [1]. This method works properly only when such region has no overlapping absorption peaks (no cross-band).

In our method we take into account the full spectral information the mixture of spectral components and can be split into more than two, without having an access to pure spectra. For example, a partially dehydrated glucose can dehydrate slightly in a presence of other evolving spectra, and a resolution of concentration profiles and pure spectral components will still be possible [4]. We consider a linear mixture model - the resulting spectrum Y_j in each moment t_j is a mixture C_j of pure spectral components S_j (with full additivity and non-negativity constraint) and an additive Gaussian noise E_j

$$Y = CS + E. \tag{1}$$

---

[*] Leading author: lukasz.sterczewski@pwr.edu.pl

Matrix *Y* contains the evolution of the THz spectrum y_j in *L* bands caused by dehydration. Dehydration process meets the full-additivity constraint (concentration of the dehydrated and hydrated components sum to one). By applying an algorithm described in [4], we can estimate the pure spectral components S and a concentration vector c_i in each observation *i*. This is done by sampling a probability density function (PDF) of the algorithm, given by

$$f(C, S, \sigma_e^2, \alpha|Y) \propto \prod_{i=1}^{N} \left[\frac{\mathbf{1}_{\mathbb{S}}(\boldsymbol{c_i})}{\sigma_{e,i}^{L+2}} \right] \exp\left(-\frac{\|\boldsymbol{y}_i - S^T \boldsymbol{c_i}\|^2}{2\sigma_{e,i}} \right) \times \prod_{m=1}^{2} \left[\frac{\Gamma(L\alpha_m + \alpha_\beta + 1)}{\left(\sum_{j=1}^{L} s_{m,j} + \beta_{\beta_m}\right)^{L\alpha_j + \alpha_{\beta_m} + 1}} \right] \times$$

$$\prod_{m=1}^{2} \left[\left(\prod_{j=1}^{L} \frac{s_{m,j}}{\Gamma(\alpha_m)} \right)^{\alpha_m - 1} \mathbf{1}_{\mathbb{R}_+^L}(\boldsymbol{s_m}) \right]. \tag{2}$$

3. Results

Dehydration of the molecule is shown in Fig. 1a. Monohydrate form exhibits absorption peaks at 59,7 and 65,5 cm^{-1}, which dehydrates into the anhydrous form with the absorption peaks at 47,7 cm^{-1}, 42,5 cm^{-1}. and 69 cm^{-1}.
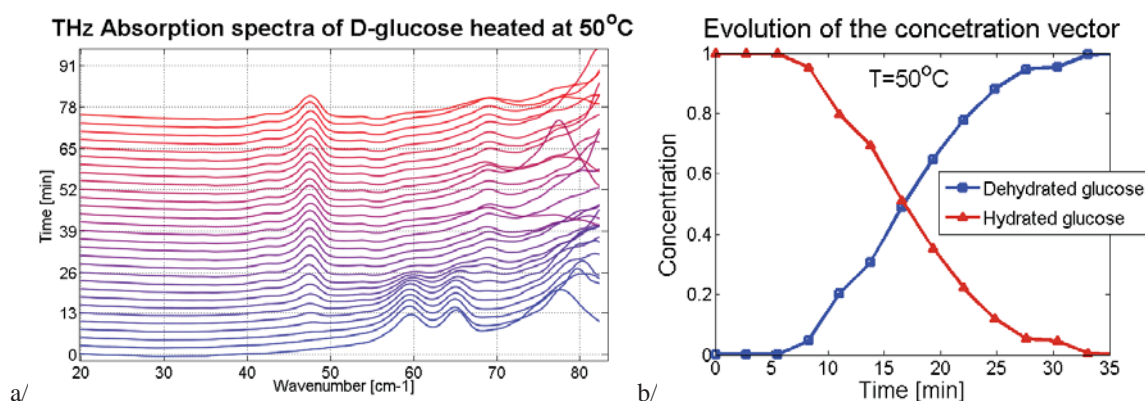

a/   b/

Figure 1: Dehydration of D-glucose monohydrate, offset of 0.2 in absorbance is introduced (a)
and a conentration profile evolution estimated by the algorithm (b)

Fig. 1b shows the concentration of glucose in hydrated in dehydrated form versus time. A total dehydration at 50°C occurs after 35 minutes. Applied algorithm thanks to evident difference in absorption peaks was capable of extracting unmixed spectra and evolution of the profile. Different profiles for each temperature can be used to find the kinetics of dehydration.

4. Conclusion

Quantitative analysis of the dehydrate can be evaluated by applying the spectral sources unmixing algorithm, which is insensitive to overlapping spectral bands in contrast with the peak area method. Estimated kinetics of dehydration plays a very important role in a drug formulation, thus its precise determination is strongly desired.

5. Acknowledgements

6. References

[1] X.-C. Zhang and H.-B. Liu, "Dehydration kinetics of D-glucose monohydrate studied using THz time-domain spectroscopy," *Chemical Physics Letters, 29(1-3),* pp. 229-233, 2006.

[2] S. L. Dexheimer, "Pharmaceutical and Security Applications of Terahertz Spectroscopy," in *Terahertz Spectroscopy, Principles and Applications*, CRC Press, 2008, pp. 299-323.

[3] H. Namkung, J. Kim, H. Chung and M. A. Arnold, "Impact of pellet thickness on quantitative terahertz spectroscopy of solid samples in a polyethylene matrix," *Analytical Chemistry, 85(7),* pp. 3674-3681, 2013.

[4] N. Dobigeon, S. Moussaoui, J.-Y. Tourneret and C. Carteret, "Bayesian separation of spectral sources under non-negativity and full additivity constraints," *Signal Processing, vol. 89 (12),* pp. 2657-2669, 2009.