Mathematical and Statistical Tools for
Online NMR Spectroscopy in Chemical Processes

Simon Kern, Svetlana Guhl, Klas Meyer, Lukas Wander, Andrea Paul, and Michael Maiwald
Bundesanstalt für Materialforschung und -prüfung (BAM), Berlin, Deutschland
Richard-Willstätter-Str. 11, D-12489 Berlin, Deutschland
michael.maiwald@bam.de

Abstract

Monitoring chemical reactions is the key to chemical process control. Today, mainly optical online methods are applied, which require excessive calibration effort. NMR spectroscopy has a high potential for direct loop process control while exhibiting short set-up times. Compact NMR instruments make NMR spectroscopy accessible in industrial and harsh environments for advanced process monitoring and control, as demonstrated within the European Union's Horizon 2020 project CONSENS.
We present a range of approaches for the automated spectra analysis moving from conventional multivariate statistical approach, (i.e., *Partial Least Squares Regression*) to physically motivated spectral models (i.e., *Indirect Hard Modelling* and *Quantum Mechanical* calculations). By using the benefits of traditional qNMR experiments data analysis models can meet the demands of the PAT community (*Process Analytical Technology*) regarding low calibration effort/calibration free methods, fast adaptions for new reactants or derivatives and robust automation schemes.

Keywords: Online NMR Spectroscopy; Process Control; Partial Least Squares Regression; Indirect Hard Modelling; Quantum Mechanics; First Principles.

Integrated Process Design – Need for Smart Field Devices

Novel concepts in the field of process engineering and in particular process intensification are currently promoted for analysis and design of innovative equipment and processing methods [1, 2]. This leads to substantially improved sustainability, efficiency, environmental performance, and alternative energy conversion.

Continuous Pharmaceutical Reaction Step in a Modular Plant

Compared to traditional batch processes, intensified continuous production gives admittance to new and difficult to produce compounds (see reaction Fig. 1 as an example),

leads to better product uniformity, and drastically reduces the consumption of raw materials and energy. Flexible (modular) chemical plants can produce various products using the same equipment with short down-times between campaigns, and quick introduction of new products to the market. Typically, such plants have smaller scale than plants for basic chemicals in batch production but still are capable to produce kilograms to tons of specialty products each day.

Consequently, full automation is a prerequisite to realize such benefits of intensified continuous plants. In continuous flow processes, continuous, auto-mated measurements and tight closed-loop control of the product quality are mandatory. If these are not available, there is a
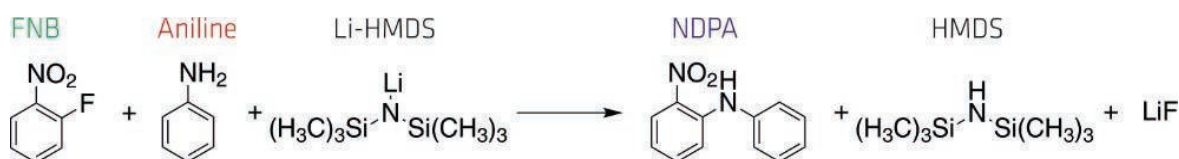

Fig. 1.   Reaction scheme: FNB: 1- fluoro-2-nitrobenzene, Li-HMDS: Lithium-bis(trimethyl-silyl)-amide, NDPA: 2-nitrodiphenylamine. Aniline was also replaced by p-toluidine and p-fluoroaniline

huge risk of producing large amounts of *Out-of-Spec* (OOS) products.

In pharmaceutical production, the common future vision is *Continuous Manufacturing* (CM), based on *Real Time Release* (RTR), i.e., a risk-based and integrated quality control in each process unit. This will allow for flexible hook-up of smaller production facilities, production transfer towards fully automated facilities, less operator intervention, less down time, and end to end process understanding over product lifecycle, future knowledge, and faster product to market. It is also assumed to significantly reduce the quality control costs within a CM concept at the same time

Fig. 1 represents a given example of a pharmaceutical reaction step, within two aromats are coupled using the lithium base Li-HMDS. The reaction takes place in a 5 wt-% solution in tetrahydrofuran. Deviations from unknown starting material and reactant concentrations together with the precipitation of LiF will lead to severe fouling and blocking of the modules. Typically, metal organic reactants are difficult to analyze due to the sensitivity to air and moisture. Thus, this example reaction was chosen in CONSENS to develop and validate a compact NMR sensor to maintain an optimal stoichiometry during the full course of the continuous production.

Smart Compact NMR Spectroscopy in Process Control

Monitoring specific information (such as physico-chemical properties, chemical reactions, etc.) is the key to chemical process control. The challenge within the project and its given lithiation reaction was to integrate a commercially available low-field NMR spectrometer [1] from a laboratory application to the full requirements of an automated chemical production environment including robust evaluation of NMR spectral data.
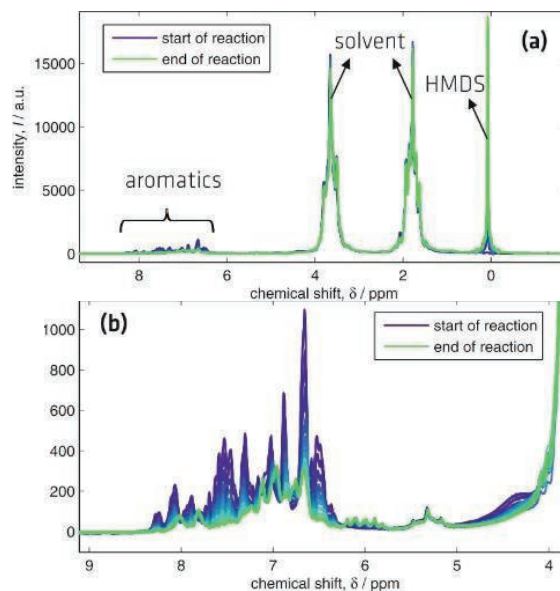

Fig. 2: (*a*) *Complete low-field NMR spectrum* (*43.5 MHz, single scan*) *and* (*b*) *aromatic region of automatically phased, baseline corrected, and shift corrected proton spectra for the lithiation reaction.*

The NMR analyzer (Fig. 3, orange box) is provided in an explosion proof housing of 57×57×85 cm module size and involves a compact 43.5 MHz NMR spectrometer togeth-
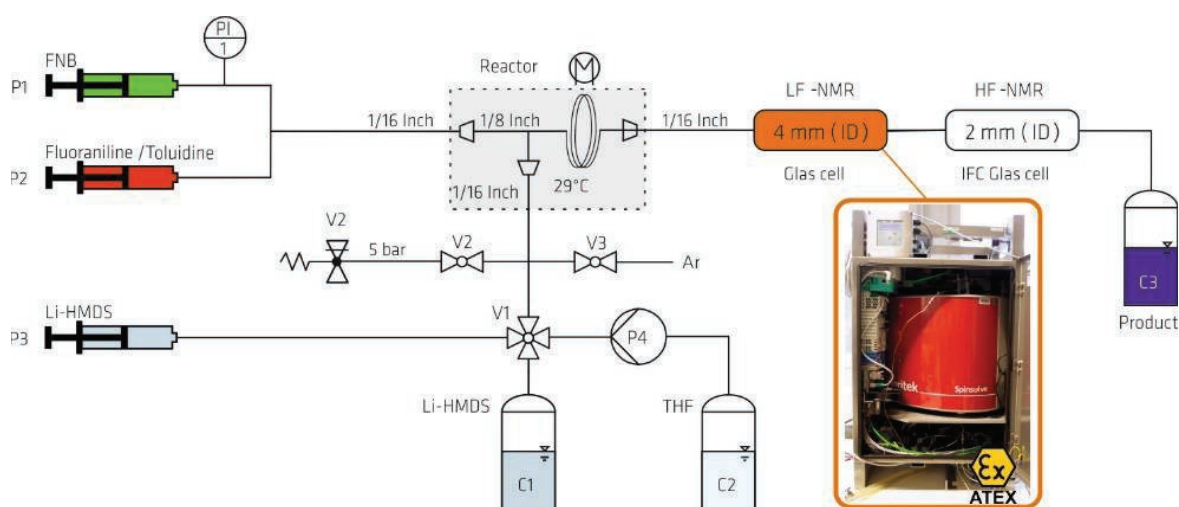

Fig. 3. *Scheme of the validation set-up for monitoring of the continuous reaction unit with the compact NMR sensor (orange box). The lithiation reaction (Fig. 1) is continuously carried out in a thermostated 1/8" tubular reactor using syringe pumps. HF NMR spectroscopy (upper right) served as reference.*

er with an acquisition unit and a programmable logic controller for automated data preparation (phasing, baseline correction) as well as data evaluation (see below). Therefore, the aromatic region of the NMR spectra in Fig. 2 had to be chosen; representing higher order NMR spectra.

In a first approach, a couple of semi batch reactions were performed for development of *Partial Least Squares Regression* (PLS-R) as well as *Indirect Hard Modeling* (IHM) models. Within these studies Li-HMDS was dosed stepwise to the reactants aniline and FNB in a batch reactor to produce spectral data material along the reaction coordinate. The reaction was in parallel observed using 500 MHz high-field NMR spectroscopy as reference method.

For validation purposes, the set-up depicted in Fig. 3 was used for monitoring of the continuous lithiation reaction in a thermostated 1/8" tubular reactor using syringe pumps. The set-up was matched to the reaction conditions of the actual plant. It was used to validate the PLS-R and IHM models as described later – again using high-field NMR spectroscopy as reference method. A considerable number of continuous experiments were performed for validation taking account for various reaction conditions by individually adjusting the flow rates of the reagents aniline, FNB, and Li-HMDS.

Data Analysis Methods

Die Chemometrics for the derivation of empirical models, e.g., PLS-R or PCA (*Principal Component Analysis*) is available and state of the art in reaction and process monitoring. Automated applications along the life cycle are still very limited. Up to now the development of such models requires significant experimental work, i.e., producing data from several time-consuming calibration runs, ideally via experimental plans (DoE).

Physically Motivated Spectral Models

The use of so-called *First Principles* methods along with reduction of the effort needed for these experiments is focus of ongoing research. Making use of novel sensors, like online NMR, in combination with flexible data analysis methods like IHM tremendously promote the use of novel process control concepts [1–3].

IHM model development consists of three steps: Firstly, *Pure Component Models* are built upon NMR spectra of the reactants and products (Fig. 4a). Each pure component model (Hard Model) consists of several Lorentzian-Gaussian functions, representing the spectral peaks [3].

Within that *Hard Model*, the ratio of peak areas are fixed against each other. Secondly, an experimental NMR spectrum is acquired and prepared by phasing and baseline correction (Fig. 4b). Finally, this experimental spectrum is
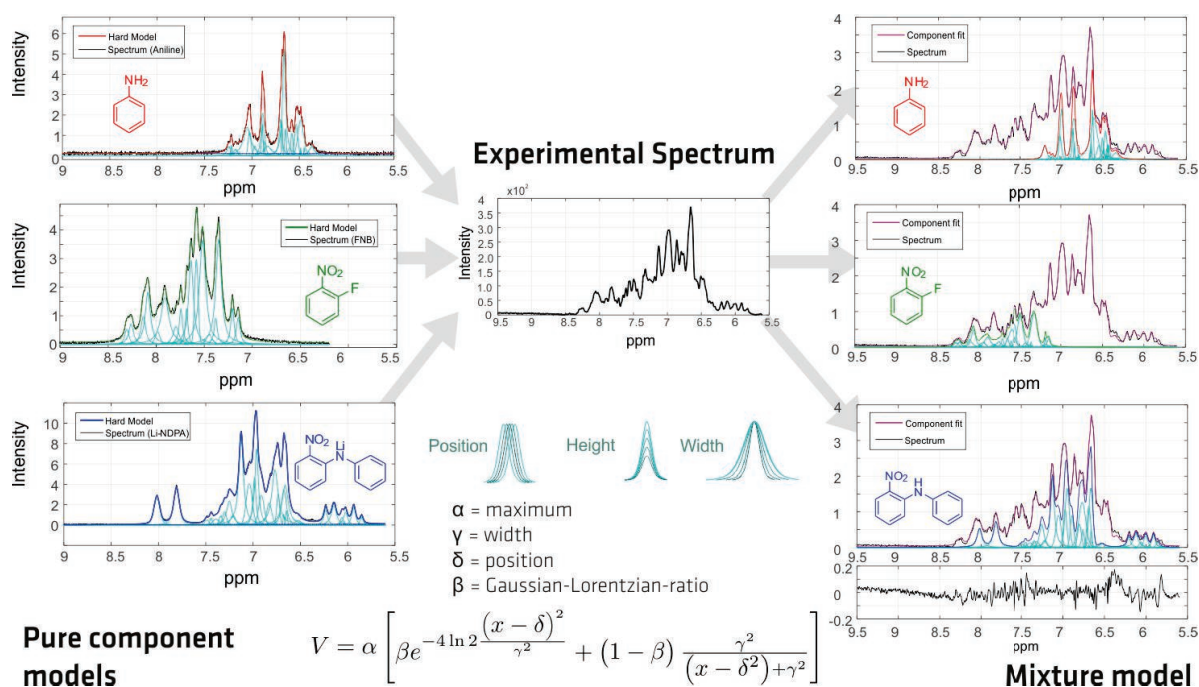

$$V = \alpha \left[\beta e^{-4\ln 2 \frac{(x-\delta)^2}{\gamma^2}} + (1-\beta) \frac{\gamma^2}{(x-\delta^2)+\gamma^2} \right]$$

Fig. 4. *Data analysis scheme for* Indirect Hard Modeling (*IHM*) *of the aromatic region* (*see Fig. 2*) *of the NMR spectrum.*

represented by the given pure component models (*Component Fitting*) from the beginning step by iterative fit routines aiming at minimized residues (Fig. 4c). Within this step, selected parameters of Lorentzian-Gaussian functions such as position, height, or width can be optimized. This allows for slight line shifts or other non-linear effects along the course of the reaction, which likely occur in NMR spectra of technical mixtures and make IHM the method of choice.
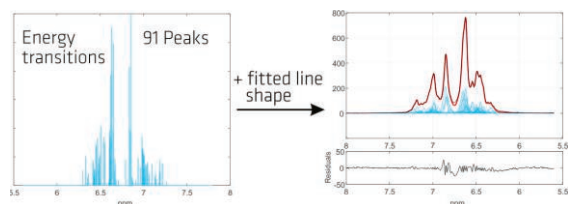

Fig. 5.   Calculation of pure component model based on spin calculations (NMR Solutions, Kuopio).

Fig. 5 depicts how physically motivated spectral models based on quantum mechanical calculations can be used to derive the pure component models, shown for aniline. Therefore, line spectra from spin calculations were adopted by empirically fitting their line shape to the real pure component spectrum.

Validation Results

Figure 6 shows the amount of substance fractions observed with the LF NMR sensor and the IHM methods according to a Design of Experiments (DoE) over an observation period of 6.5 hours in comparison with HF NMR data. In some runs aniline was replaced by p-toluidine or p-fluoroaniline (Fig. 1, 3) in order to test the modular IHM approach. In all cases, the pure component models were exchangeable and worked together with the remaining

models for FNB. In general, all results found by IHM are in good agreement with results from the PLS-R model as well as the HF NMR reference data.

The largest prediction uncertainties for IHM were found for aniline to be 5–7 mol-% relative (aromats), i.e., 0.25–0.35 mol-% absolute, or 20–30 mmol L^{-1} concentration deviation. As can be seen in Figure 4a, the signals of aniline completely overlap with the further reagents causing such model deviations. IHM slightly overestimates aniline in the low concentration range during equivalent fitting of the three pure component models, thus, underestimating the product NDPA. Minimizing the residues presents a closing condition for the fitting process. Improving IHM for these unwanted issues is currently undertaken.

References

[1] K. Meyer, S. Kern, N. Zientek, G. Guthausen and M. Maiwald, Process Control with Compact NMR, Trends in Analytical Chemistry 83, 39–52 (2016); doi: 10.1016/j.trac.2016.03.016

[2] M. Maiwald, P. Gräßer, L. Wander, N. Zientek, S. Guhl, K. Meyer and S. Kern, Strangers in the Night—Smart Process Sensors in Our Current Automation Landscape, Proceedings 1, 628 (2017); doi: 10.3390/proceedings1040628

[3] A. Michalik-Onichimowska, S. Kern, J. Riedel, U. Panne, R. King and M. Maiwald, "Click" analytics for "click" chemistry – A simple method for calibration-free evaluation of online NMR spectra, Journal of Magnetic Resonance 277, 154–161 (2017); doi: 10.1016/j.jmr.2017.02.018
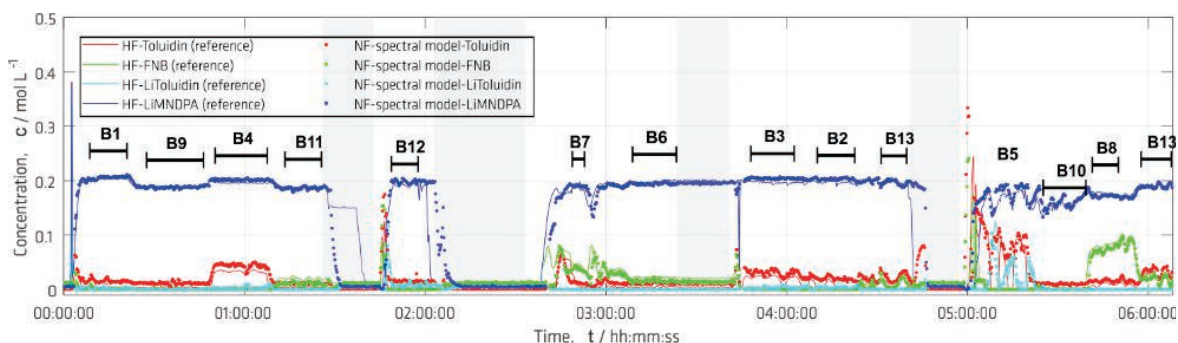
Fig. 6.   Amount of substance fractions observed with the LF NMR sensor of the reagents aniline and FNB and the product NDPA along an observation period of 6.5 hours together with HF NMR data. Grey areas represent points in time where pumps were not running due to cleaning or refilling.