On the uniqueness of continuous and discrete hard models of NMR-spectra

Jan Hellwig[a,b], Klaus Neymeyr[a,b]

[a] *Universität Rostock, Institut für Mathematik, 18051 Rostock, Germany*
[b] *Leibniz-Institut für Katalyse e.V., 18059 Rostock, Germany*

Abstract

Lorentz, Gauss, Voigt and pseudo-Voigt functions play an important role in hard modeling of NMR spectra. This paper shows the uniqueness of continuous NMR hard models in terms of these functions by proving their linear independence. For the case of discrete hard models, where the spectra are represented by finite-dimensional vectors, criteria are given under which the models are also unique.

Keywords: NMR, hard modeling, model uniqueness, pseudo-Voigt functions, Voigt-functions

1. Introduction

1.1. Overview

NMR spectroscopy is an analytical technique for elucidating the structure of chemical compounds by extracting qualitative and quantitative information from the spectra. *Hard modeling* [2] is a first step in the quantitative analysis of NMR spectra. It attempts to reproduce a given spectrum as a sum of individually parametrized model functions. The model parameters determine the chemical and physical structure of the system for which the NMR spectra are measured.

Hard models of NMR spectra typically use model functions of the Lorentz, pseudo-Voigt or Voigt type [12, 6, 8, 15, 1]. The goal is to find an optimal fit of a given spectrum in terms of a finite dimensional expansion of such profiles. The task of assigning an appropriate number of peak profiles and determining their parameters can be formulated as a (sometimes high-dimensional) optimization problem. This problem can be solved by numerical optimization [2, 9] or by neural networks [7, 16]. The model fitting approaches raise the question of whether the hard models contain some model or parameter ambiguity. Here, we show that such hard models are unique from a rigorous mathematical point of view. Comparable analyses of linear independent and even orthogonal sets of base functions have already been presented in the area of so-called Gabor systems [3] and further, wavelet theory [4], but we are not aware of any analyses of a hard model ambiguity in terms of NMR peak model functions.

The uniqueness question can no longer be answered positively if only approximate expansions in terms of the model functions are considered or if the data are only given up to an accuracy threshold. Such a situation exists, for example, for noisy experimental data. In this case, the uniqueness property is increasingly lost as the error tolerance level increases. An example is presented in Appendix C.1.

This work is organized as follows. First, we give a strict mathematical form to the hard modeling problem . Then, in Section 2, we prove the linear independence of finite sets of differently parametrized model functions and thus show the continuous model uniqueness. Section 3 discusses discrete hard models and gives criteria under which discrete hard models are also unique.

1.2. Problem definition

Lorentz, Gauss, Voigt and pseudo-Voigt profiles are introduced first. A Lorentzian with the center value c and the half-width parameter $w > 0$ corresponds to an ideal NMR signal without noise or other physical interferences

$$L_{c,w}(x) = \frac{w^2}{w^2 + (x - c)^2}. \tag{1}$$

A Gaussian with parameters $(c, w) \in \mathbb{R} \times \mathbb{R}_{>0}$ is defined by

$$G_{c,w}(x) = \exp\left(-\ln(2) \cdot \frac{(x-c)^2}{w^2}\right) \tag{2}$$

and can compensate for such interferences when combined with a Lorentzian, broadening the range of line shapes.

A Voigt profile with parameters $(c, v, w) \in \mathbb{R} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ is given by

$$V_{c,v,w}(x) = (G_{0,v} * L_{c,w})(x) = \int_{-\infty}^{\infty} G_{0,v}(y) \cdot L_{c,w}(x-y)\mathrm{d}y \tag{3}$$

and consists of a Lorentzian convolved with a centered Gaussian at possibly different half widths. Allowing the Gaussian to be uncentered does not enable a more general type of function, refer to Appendix C.3.

A pseudo-Voigt function, also called Gauss-Lorentz function, with parameters $(c, w, \lambda) \in \mathbb{R}\times\mathbb{R}_{>0}\times[0, 1]$ simplifies Voigtian line shape calculations through the use of a convex combination of Gaussian and Lorentz functions with identical parameters

$$PV_{c,w,\lambda}(x) = \lambda \cdot G_{c,w}(x) + (1-\lambda) \cdot L_{c,w}(x). \tag{4}$$

Next, we refer to the different peak model functions in terms of the variable f_p, where p is a parameter tuple (e.g $p = (c, w)$). We assume that for all p the function f_p is absolutely integrable and non-zero, i.e. $f_p \in L^1(\mathbb{R}) \setminus \{0\}$; this is required for the Fourier transform. Additionally, we define $[n]$ as the set $\{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}_{>0}$.

Definition 1.1. *An* NMR spectrum *is a function $s : \mathbb{R} \to \mathbb{R}$ represented by a sum of a positive finite number of model functions, each parametrized by pairwise distinct tuples p_i and scaling coefficients $\lambda_i \in \mathbb{R} \setminus \{0\}$, where*

$$s(x) := \sum_{i=1}^{n} \lambda_i \cdot f_{p_i}(x). \tag{5}$$

The set $\{(\lambda_i, p_i) \mid i \in [n]\}$ is called a hard model *or* decomposition *of s.*

Remark 1.2. *We can exclude identical parameter tuples, since f_{p_i} and f_{p_j} can be combined into $(\lambda_i + \lambda_j)f_{p_i}(x)$, if $p_i = p_j$ for $i \neq j$, and we can effectively remove one summand. Removing redundant functions from the spectra definition eliminates the possibility of mathematically trivial ambiguities of the type $f_p = \sum_{i=1}^{k} \lambda_i f_p$, when $\sum_{i=1}^{k} \lambda_i = 1$. In terms of chemistry, this means assigning multiple models of identical widths and positions to a single peak. Usually, the model functions $f_{p_i}(x)$ and scaling parameters λ_i are non-negative, implying $s(x) \geq 0$. In Section 2, we examine the situation in which $\lambda_i < 0$ are permitted.*

Definition 1.3. *A hard model $\{(\lambda_i, p_i) \mid i \in [n]\}$ of an NMR spectrum s is called* ambiguous*, if another hard model $\{(\mu_j, q_j) \mid j \in [m]\}$ of s exists for some $m \in \mathbb{N}_{>0}$ and $\mu_j \in \mathbb{R} \setminus \{0\}$. Therefore an ambiguously decomposable spectrum s can be represented as*

$$s(x) = \sum_{i=1}^{n} \lambda_i f_{p_i}(x) = \sum_{j=1}^{m} \mu_j f_{q_j}(x). \tag{6}$$

If a hard model is not ambiguous, then it is called unique.

Furthermore, we call a pair of ambiguous hard models $\{(\lambda_i, p_i) \mid i \in [n]\}$ and $\{(\mu_j, q_j) \mid j \in [m]\}$ minimal*, if $p_i \neq q_j$ holds for all i, j.*

If there are indices k, l in the settings of Definition 1.3 satisfying $p_k = q_l$, spectra with minimal ambiguous hard models can be constructed. This concept is illustrated in Appendix C.2. In cases where these indices exist, we can subtract $\min\{\lambda_k, \mu_l\} \cdot f_{p_k}(x)$ from Eq. (6), resulting in another spectrum $s'(x)$ with an ambiguous decomposition, and one or two fewer summands. If the resulting hard model is not minimal, we can repeat the procedure. Eventually, this results in one of three cases.

1) The resulting two sums have at least one summand and are therefore minimally ambiguous.
2) One sum vanishes while the other one does not. Given that f is neither the zero function nor does the scaling coefficient vanish, the other sum must have at least two summands of different parameter tuples. We can subtract the first remaining summand to obtain a spectrum with minimally ambiguous hard models.

3) Both sums vanish, which implies that the summands of the initial hard model are merely permutated. This implies that the initial decomposition sets are equal and contradict the definition.

So the existence of minimally ambiguous hard models is equivalent to the existence of ambiguous hard models and it suffices to show that no minimally ambiguous hard models exist for certain model functions f.

Additionally, minimal uniqueness is not primarily a property of a particular spectrum, but a property of the peak model functions and their parametrizations. We can set $k = n + m$ and $\lambda_{n+j} = -\mu_j, p_{n+j} = q_j$ for $j \in [m]$ and rewrite Eq. (6) as follows

$$0 = \sum_{i=1}^{k} \lambda_i f_{p_i}(x). \tag{7}$$

If there is an ambiguous hard model, then there is some $k \in \mathbb{N}$ and parameter tuples p_i, such that the set $\{f_{p_i} \mid i \in [k]\}$ is linearly dependent. Conversely, if $k \geq 2$, $\lambda_i \in \mathbb{R} \setminus \{0\}$ and parameter tuples p_i exist, such that Eq. (7) holds, we can construct spectra which have ambiguous hard models, namely

$$s(x) = \lambda_1 f_{p_1}(x) = \sum_{j=2}^{k} -\lambda_j f_{p_j}(x).$$

Therefore, the independence of a finite number of model functions with different parameters is equivalent to the uniqueness of hard models of all NMR spectra belonging to this type of model function. Finally, we always consider linear independence over the space of functions $f : \mathbb{R} \to \mathbb{R}$ before the Fourier transform and over the space of functions $f : \mathbb{R} \to \mathbb{C}$ after the Fourier transform.

2. Unique Hard Models of Continuous Spectra

2.1. Linear Independence of Lorentz Functions

Theorem 1. *Any finite number of pairwise different Lorentzians L_{c_j,w_j} is linearly independent.*

This theorem can be proven in two ways. The first method for proving the theorem is straightforward and uses the structure of the polynomials that underlie the Lorentz functions. The second proof uses the Fourier transformation and works with technical steps that lay the foundation for later, more complex proofs.

Proof by algebraic means. We assume a linear combination of pairwise different Lorentzians to be given with coefficients $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

$$0 = \sum_{j=1}^{n} \lambda_i \frac{w_j^2}{w_j^2 + (x - c_j)^2}$$

with pairwise different parameter tuples $p_i = (c_j, w_j) \in \mathbb{R} \times \mathbb{R}_{>0}$. We need to show $\lambda_j = 0$ for all j to prove the linear independence. By multiplication with all denominators we get

$$0 = \sum_{j=1}^{n} \lambda_j w_j^2 \underbrace{\left(\prod_{\substack{k=1 \\ k \neq j}}^{n} w_k^2 + (x - c_k)^2 \right)}_{:= g_j(x)}. \tag{8}$$

For any j the polynomial g_j has the degree $2n - 2$. Its set of zeros $\{c_k \pm w_k \mathbf{i} \mid k \neq j\}$ consists of $2n - 2$ pairwise different complex numbers (since $w_k > 0$) and there are no other zeros. Thus $z_j := c_j + w_j \mathbf{i}$ cannot be a zero of g_j, i.e. $g_j(z_j) \neq 0$. Inserting $z_k = c_k + w_k \mathbf{i}$ in (8) for $k = 1, \ldots, n$ results in

$$0 = \sum_{j=1}^{n} \lambda_j w_j^2 g_j(z_k) = \lambda_k \underbrace{w_k^2}_{>0} \underbrace{g_k(z_k)}_{\neq 0}.$$

Thus $\lambda_k = 0$ for all k. This proves that any set of finitely many pairwise different Lorentz functions is linearly independent. □

Before proving Thm. 1 in another way, let us define the Fourier transform.

Definition 2.1. *Let $f \in L^1(\mathbb{R})$. The continuous Fourier transform $\hat{f}$ is defined by*

$$\hat{f}(\omega) = (\mathcal{F} f)(\omega) = \int_{\mathbb{R}} f(x) \mathrm{e}^{-\mathbf{i}x\omega} \mathrm{d}x, \qquad \omega \in \mathbb{R}.$$

Since the Fourier transform is linear and invertible, if f is continuous [11, Thms. 2.8, 2.10], it is straightforward to show that a set of functions $\{f_i, \mid i \in [n]\}$ is linearly independent if and only if the set of Fourier transformed functions $\{\hat{f}_i \mid i \in [n]\}$ is also linearly independent. Thus, proving Theorem 1 only requires showing the linear independence of the Fourier transformed functions.

Lemma 2.2. *Assume $f = L_{c,w}$ with parameters $c \in \mathbb{R}$ and $w \in \mathbb{R}_{>0}$. Then, for $\omega \in \mathbb{R}$ we have*

$$(\mathcal{F} f)(\omega) = C(w) \cdot \mathrm{e}^{-w|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c\omega},$$

where $C(w) \in \mathbb{R}$ is a constant only dependent on the parameter w of f.

For its proof, see Appendix A.

The Fourier transformed Lorentzians converge at significantly different rates for different w_j. The dominating term is defined by the functions with the smallest value of w_j. We can eliminate all functions with greater convergence rates, resulting in a simplified case, where only the parameters c_j remain. We can prove that the convergence of the remaining sum of functions implies that all of the remaining coefficients must equal 0.

Lemma 2.3. *For $j \in [m]$ let $c_j \in \mathbb{R}$ be pairwise distinct and let $\lambda_j \in \mathbb{C}$ such that*

$$\sum_{j=1}^{m} \lambda_j \mathrm{e}^{-\mathbf{i}c_j\omega} \overset{\omega\to\infty}{\to} 0. \tag{9}$$

Then, all $\lambda_j = 0$.

Intuitively, a sum of periodic functions can hardly converge, if it is not already constant. The proof of this lemma only involves the use of basic mathematical theorems. Therefore, it is omitted here and can be found in Appendix B.1.

The following proof is conducted as a proof by contradiction.

Proof of Thm. 1 by means of Fourier transformations. Let $\{(c_j, w_j) \mid j \in [n]\} \subseteq \mathbb{R} \times \mathbb{R}_{>0}$ be a set of distinct ordered pairs. We assume that coefficients $\lambda_j \in \mathbb{C} \setminus \{0\}$ exist such that

$$\sum_{j=1}^{n} \lambda_j \mathrm{e}^{-w_j|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c_j\omega} = 0. \tag{10}$$

Additionally, we define $w = \min\{w_j \mid j \in [n]\}$ and $I = \{j \in [n] \mid w_j = w\} \neq \emptyset$. We then multiply Eq. (10) by $\mathrm{e}^{w|\omega|}$, thereby dividing by the smallest rate of decay

$$\sum_{j \in I} \lambda_j \mathrm{e}^{-\mathbf{i}c_j\omega} + \sum_{j \in [n]\setminus I} \lambda_j \mathrm{e}^{-(w_j - w)|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c_j\omega} = 0.$$

As ω approaches infinity, the functions with greater rates of decay converge to 0 and the above equation implies

$$\sum_{j \in I} \lambda_j \mathrm{e}^{-\mathbf{i}c_j\omega} \overset{\omega\to\infty}{\to} 0. \tag{11}$$

Note that for $j \in I$ we have $w_j = w$ and therefore all c_j are distinct. Finally, Lemma 2.3 implies $\lambda_j = 0$ for $j \in I$, thereby contradicting the initial assumption. □

2.2. Linear Independence of Gaussian Functions

Theorem 2. *Any finite number of pairwise different Gaussians G_{c_j,w_j} is linearly independent.*

A well-known result states that applying the Fourier transformation to a Gaussian function again yields a Gaussian function [11, Ex. 2.6]. While it is possible to prove the independence of multiple Gaussians without Fourier transforming them, applying the Fourier transformation first allows for a consistent method of proof for all four model functions (1) - (4). In addition, the Fourier transformed Gaussians are also required in Section 2.3.

Lemma 2.4. *Assume $f = G_{c,w}$ with parameters $c \in \mathbb{R}$ and $w \in \mathbb{R}_{>0}$. Then, for $\omega \in \mathbb{R}$, we have*

$$(\mathcal{F} f)(\omega) = C(w) \cdot \mathrm{e}^{-w'^2\omega^2} \cdot \mathrm{e}^{-\mathbf{i}c\omega},$$

where $C(w) \in \mathbb{C}$ is a constant only dependent on the parameter w of f and $w' = \frac{w}{2\sqrt{\ln 2}}$.

For the proof, see Appendix A.

Proof of Thm. 2. It suffices to show that an arbitrary, finite set of Fourier transformed Gaussians is linearly independent. Thus, we can neglect the difference between w' and w in this proof. We conduct this proof similarly to the second proof of Thm. 1. Assume that $\lambda_j \in \mathbb{C} \setminus \{0\}$ and $(c_j, w_j) \in \mathbb{R} \times \mathbb{R}_{>0}$ exist such that

$$\sum_{j=1}^{n} \lambda_j \mathrm{e}^{-w_j^2\omega^2} \cdot \mathrm{e}^{-\mathbf{i}c_j\omega} = 0. \tag{12}$$

We define $w = \min_{j\in[n]} w_j$ and $I = \{j \in [n] \,|\, w_j = w\}$. Multiplying Eq. (12) by $\mathrm{e}^{w^2\omega^2}$ and taking the limit as ω approaches infinity yields

$$\sum_{j\in I} \lambda_j \mathrm{e}^{-\mathbf{i}c_j\omega} + \sum_{j\in[n]\setminus I} \lambda_j \mathrm{e}^{-(w_j^2-w^2)\omega^2} \cdot \mathrm{e}^{-\mathbf{i}c_j\omega} \overset{\omega\to\infty}{\to} 0.$$

The second sum converges to 0 and we have

$$\sum_{j\in I} \lambda_j \mathrm{e}^{-\mathbf{i}c_j\omega} \overset{\omega\to\infty}{\to} 0.$$

This is the same as Eq. (11). Lemma 2.3 implies that $\lambda_j = 0$ for $j \in I \neq \emptyset$, contradicting the assumption. □

2.3. Linear Independence of Voigt Functions

Theorem 3. *Any finite number of pairwise different Voigt functions V_{c_j,v_j,w_j} is linearly independent.*

Let us recall a well-known fact that helps proving Theorem 3.

Lemma 2.5. *Let $f, g \in L^1(\mathbb{R})$. Then for all $\omega \in \mathbb{R}$*

$$(\mathcal{F}(f * g))(\omega) = (\mathcal{F} f)(\omega) \cdot (\mathcal{F} g)(\omega) \tag{13}$$

holds.

This lemma can be found and is proven in [11, Thm. 2.15].

Lemma 2.6. *Assume $f = V_{c,v,w}$ with parameters $c \in \mathbb{R}$ and $v, w \in \mathbb{R}_{>0}$. Then, for $\omega \in \mathbb{R}$, we have*

$$(\mathcal{F} f)(\omega) = C(v, w) \cdot \mathrm{e}^{-v'^2\omega^2} \cdot \mathrm{e}^{-w|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c\omega},$$

where $C(v, w) \in \mathbb{C}$ is a constant only dependent on the parameters v, w of f and $v' = \frac{v}{2\sqrt{\ln 2}}$.

Proof. Using Lemmas 2.2, 2.4 and 2.5, it follows that

$$(\mathcal{F}V_{c,v,w})(\omega) = (\mathcal{F}(G_{0,v} * L_{c,w}))(\omega) \overset{2.5}{=} (\mathcal{F}G_{0,v})(\omega) \cdot (\mathcal{F}L_{c,w})(\omega)$$
$$\overset{2.2,2.4}{=} C(v,w) \cdot \left(\cdot \mathrm{e}^{-v'^2\omega^2}\right) \cdot \left(\mathrm{e}^{-w|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c\omega}\right).$$

□

Proof of Thm. 3. Again, we may neglect the difference of v' and v for this proof. We assume there are pairwise different triples $(c_j, v_j, w_j) \in \mathbb{R} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ and $\lambda_j \in \mathbb{C} \setminus \{0\}$ such that

$$0 = \sum_{j=1}^{n} \lambda_j \underbrace{\mathrm{e}^{-v_j^2\omega^2} \cdot \mathrm{e}^{-w_j|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c_j\omega}}_{:= g_j(\omega)}. \tag{14}$$

Additionally, we define

$$v = \min\{v_j^2 \mid j \in [n]\}, \quad w = \min\{w_j \mid v_j^2 = v, j \in [n]\}$$
$$I_1 = \{j \in [n] \mid v_j^2 = v, w_j = w\} \neq \emptyset, I_2 = \{j \in [n] \mid v_j^2 = v, w_j \neq w\}, I_3 = [n] \setminus (I_1 \cup I_2).$$

All functions g_j with $j \in I$ have the smallest rate of convergence. Multiplying Eq. (14) by $\mathrm{e}^{v\omega^2 + w|\omega|}$ leads to

$$0 = \sum_{j\in I_1} \lambda_j \mathrm{e}^{-\mathbf{i}c_j\omega} + \sum_{j\in I_2} \lambda_j \mathrm{e}^{-(w_j-w)|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c_j\omega} + \sum_{j\in I_3} \lambda_j \mathrm{e}^{-(v_j^2-v)\omega^2} \cdot \mathrm{e}^{-(w_j-w)|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c_j\omega}.$$

As ω approaches infinity, the second and the third sum vanish, because for $j \in I_2$ the inequality $w_j > w$ holds, and for $j \in I_3$ we have $v_j > v$ and $e^{-a\omega^2 + b\omega} \to 0$ for all $a > 0, b \in \mathbb{R}$. Consequently, we have the following convergence

$$\sum_{j\in I_1} \lambda_j \mathrm{e}^{-\mathbf{i}c_j\omega} \overset{\omega\to\infty}{\to} 0.$$

This is identical to Eq. (11). Finally, Lemma 2.3 implies $\lambda_j = 0$ for $j \in I_1$ and $I_1 \neq \emptyset$, leading to a contradiction. □

2.4. Linear Independence of Pseudo-Voigt Functions

The use of pseudo-Voigt functions introduces another type of ambiguity linked to the parameter $\lambda \in [0, 1]$. It is important to note this ambiguity when considering the linear independence of the functions.

Remark 2.7. *Let $n > 1$, $\lambda_1, \ldots, \lambda_n \in [0, 1]$, $\alpha_1, \ldots, \alpha_n \in \mathbb{R}_{>0}$, and $(c, w) \in \mathbb{R} \times \mathbb{R}_{>0}$. With*

$$\alpha' = \sum_{i=1}^{n} \alpha_i, \qquad \lambda' = \frac{\sum_{i=1}^{n} \alpha_i \lambda_i}{\alpha'},$$

we have

$$\sum_{i=1}^{n} \alpha_i PV_{c,w,\lambda_i} = \left(\sum_{i=1}^{n} \alpha_i \lambda_i\right) G_{c,w} + \left(\sum_{i=1}^{n} \alpha_i (1 - \lambda_i)\right) L_{c,w}$$
$$= \alpha' \lambda' G_{c,w} + \alpha'(1 - \lambda') L_{c,w} = \alpha' PV_{c,w,\lambda'}. \tag{15}$$

The above remark implies that any spectrum with two or more summands of different λ and identical parameters (c, w) can be represented by using fewer summands and therefore has an ambiguous hard model. This works due to λ behaving similarly to a linear scaling parameter. In fact, with some redefinitions the problem of fitting λ_i as well as some height parameters α_i can be solved by linear least squares.

If we consistently write such ambiguities as a single summand, we can assume that for different pseudo-Voigt functions the tuples (c, w) are distinct. We can even prove a more general result.

Theorem 4. *For $j \in [n]$, $k \in [m]$ let $p_j = (c_j, v_j)$ and $q_k = (d_k, w_k)$ be distinct parameter tuples with $v_j, w_k > 0$. However, we allow for equalities $p_j = q_k$. Then, the set of functions*

$$\left\{G_{c_j,v_j} \mid j \in [n]\right\} \cup \left\{L_{d_k,w_k} \mid k \in [m]\right\}$$

is linearly independent.

In other words, any finite linear combination of Lorentzians and Gaussians is linearly independent.

Proof of Thm. 4. Using the previous scheme, we only consider the Fourier transformed functions $f_j(\omega) = \mathrm{e}^{-v_j^2\omega^2} \cdot \mathrm{e}^{-\mathbf{i}c_j\omega}$ and $g_k(\omega) = \mathrm{e}^{-w_k|\omega|} \cdot \mathrm{e}^{-\mathbf{i}d_k\omega}$ and assume there are $\lambda_j, \mu_k \in \mathbb{C} \setminus \{0\}$, such that

$$\sum_{j=1}^{n} \lambda_j f_j + \sum_{k=1}^{m} \mu_k g_k = 0. \tag{16}$$

Using Thm. 2, we can deduce that $m \geq 1$, since a sum of Fourier transformed Gaussians can only equal zero if all $\lambda_j = 0$. Therefore, $w = \min\{w_k \mid k \in [m]\}$ exists and $I = \{k \in [m] \mid w_k = w\} \neq \emptyset$. If we multiply Eq. (16) by $e^{w|\omega|}$, we have

$$\sum_{j=1}^{n} \lambda_j \mathrm{e}^{-v_j^2\omega^2 + w|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c_j\omega} + \sum_{k \in I} \mu_k \mathrm{e}^{-\mathbf{i}d_k\omega} + \sum_{k \in [m]\setminus I} \mu_k \mathrm{e}^{-(w_k - w)|\omega|} \cdot \mathrm{e}^{-\mathbf{i}d_k\omega} = 0.$$

The first and the third sum each converge to 0 as ω approaches infinity and we arrive at Eq. (11) with differently named variables. Lemma 2.3 implies $\mu_k = 0$ for all $k \in I$ and $I \neq \emptyset$, contradicting the assumption. □

The following corollary indicates that the only type of ambiguity is the one mentioned in Remark 2.7.

Corollary 2.8. *Assume s is an NMR spectrum with ambiguous decompositions (α_j, p_j), $j \in [n]$ and (β_k, q_k), $k \in [m]$, where $p_j = (c_j, v_j, \lambda_j)$, $q_k = (d_k, w_k, \mu_k)$ and*

$$s(x) = \sum_{j=1}^{n} \alpha_j \cdot PV_{c_j,v_j,\lambda_j}(x) = \sum_{k=1}^{m} \beta_k \cdot PV_{d_k,w_k,\mu_k}(x). \tag{17}$$

Then for each $j \in [n]$ there exists a $k \in [m]$ so that $(c_j, v_j) = (d_k, w_k)$.

Proof. By applying Eqs. (15), we can modify both sides of Eq. (17) in such a way that each pair of parameters (c_j, v_j) and (d_k, w_k) appears not more than once in each sum. We call the remaining index sets J and K respectively. Now, assume there exists a $j_0 \in J$ such that $(c_{j_0}, v_{j_0}) \neq (d_k, w_k)$ for all $k \in K$. Consequently,

$$\alpha_{j_0}\lambda_{j_0} G_{c_{j_0},v_{j_0}}(x) = \alpha_{j_0}(1 - \lambda_{j_0}) L_{c_{j_0},v_{j_0}} + \sum_{j \in J\setminus\{j_0\}} -\alpha_j\lambda_j G_{c_j,v_j} - \alpha_j(1 - \lambda_j) L_{c_j,v_j} + \sum_{k \in K} \beta_k\mu_k G_{d_k,w_k} + \beta_k(1 - \mu_k) L_{d_k,w_k}$$

represents a non-trivial linear combination for the Gaussian on the left hand side, whose parameters (c_{j_0}, v_{j_0}) do not appear on any Gaussian on the right. This contradicts Thm. 4. □

In other words, NMR spectra consisting of pseudo-Voigt functions with different tuples (c_j, w_j) have a unique hard model.

3. Unique and Ambiguous Hard Models of Discrete Spectra

Experimental NMR spectra record signals as a finite sequence of discrete chemical shifts, rather than a continuous function. In this section we discuss the extent to which the uniqueness of hard models for continuous NMR spectra is still valid when the spectra are given only at a finite set of points X. Therefore, we need to redefine ambiguity in such a way that it suffices if Eq. (6) holds for all $x \in X$. Given n linearly independent functions $f_1, \ldots, f_n$ one can always choose n points $X = \{x_1, \ldots, x_n\}$, so that the matrix $A(i, j) := f_j(x_i)$ is invertible and the n equations

$$s(x_i) = \sum_{j=1}^{n} \lambda_j \cdot f_j(x_i) \qquad i \in [n]$$

have a unique solution in $\lambda_1, \ldots, \lambda_n$. Nonetheless, there might be sets of n points X at which A is non-invertible even for linearly independent functions.

In general, increasing the number of data points $k := |X|$ leads to a higher number of constraints, which are to be satisfied by the hard model as written in terms of the respective function bases. The limit case with infinitely many data points corresponds to continuous spectra and unique hard models. Another crucial quantity is the number of model functions $m, n \in \mathbb{N}$. Increasing their number provides more degrees of freedom, which tends towards ambiguous hard models due to a larger number of parameters to solve k equations. So we want to find the number of points $k_f(m, n)$ depending on the function class f, which guarantees the uniqueness of a hard model of a sum of n functions in m differently parametrized functions.

Definition 3.1. *Consider a parameter tuple p and let $f_p : \mathbb{R} \to \mathbb{R}$. Additionally, let $m, n \in \mathbb{N}$. We define $k_f(m, n)$ as the maximum $k \in \mathbb{N}$, so that a set $X \subseteq \mathbb{R}$ of k points as well as $m + n$ parametrized functions $f_{p_1}, \ldots, f_{p_n}, f_{q_1}, \ldots, f_{q_m}$ with positive scaling coefficients $\lambda_i, \mu_j \in \mathbb{R}_{>0}$ exist fulfilling*

$$\sum_{i=1}^{n} \lambda_i f_{p_i}(x) = \sum_{j=1}^{m} \mu_j f_{q_j}(x) \qquad \text{for all } x \in X.$$

Remark 3.2. *It should be noted that not all such sets X are finite. For instance, if $f_c(x) = \sin(x - c)$ we choose $c_1 = 0$, $c_2 = \pi$ and have $\sin(x) = \sin(x - \pi)$ for $x \in \{k\pi \mid k \in \mathbb{Z}\}$. In these cases, we conventionally write $k_{\sin}(1, 1) = \infty$.*

Note that in general $k_f(m, n) = k_f(n, m)$ and for $m_1 \le m_2$ and $n_1 \le n_2$ we have $k_f(m_1, n) \le k_f(m_2, n)$ and $k_f(m, n_1) \le k_f(m, n_2)$ respectively.

$k_f(m, n)$ is closely linked to the uniqueness of spectra hard models over a finite set X. If a spectrum s consisting of n functions is measured at any $k_f(m, n) + 1$ points, there is no hard model of up to m differently parametrized functions which is equal to s at all $k_f(m, n) + 1$ points.

If negative λ_i, μ_j were allowed in Def. 3.1, we would have $k_f(m, n) = k_f(m - 1, n + 1) = k_f(m + 1, n - 1)$, given that $m - 1, n - 1$ are greater than 0. This would imply that k_f depends solely on $m + n$. However, there may exist a function f where $k_f(1, 3) \neq k_f(2, 2)$, if limited to positive linear combinations.

As exact values for k_f cannot yet be obtained for all model functions (1) - (4), our aim is to at least establish lower and upper bounds.

Theorem 5. *Assume $f_{0,1} \in L^1(\mathbb{R}) \setminus \{0\}$ and let $\mathcal{F}$ be the set of functions $\left\{f_{c,w} = f_{0,1}\left(\frac{x-c}{w}\right) \,\middle|\, c \in \mathbb{R}, w \in \mathbb{R}_{>0}\right\}$.*
We impose the following set of conditions on $\mathcal{F}$:

(I) The function $f_{0,1}$ is continuous.

(II) The function $f_{0,1}$ is even, i.e. $f_{0,1}(x) = f_{0,1}(-x)$ for all $x \in \mathbb{R}$.

(III) The function $f_{0,1}$ is monotonically decreasing for $x > 0$.

(IV) For all $c \in \mathbb{R}, w > 1$ there exists some $\vartheta_w \in \left[0, \frac{1}{w}\right)$ such that

$$\frac{f_{c,1}(x)}{f_{0,w}(x)} \overset{x\to\infty}{\to} \vartheta_w.$$

(V) $f_{0,1}(0) = 1$ and $f_{0,1}(1) = \frac{1}{2}$.

Then $k_f(m, n) \ge 2m + 2n - 2$.

Given $m \geq n \in \mathbb{N}$, the proof constructs two sets of m and n functions with different parameters, whose scaled sums have exactly $2m+2n-2$ intersections. The intersection points cannot be given explicitly, but are implied by the intermediate value theorem [13], since we can explicitly name $2m+2n-1$ points where the first sum is strictly greater than the second sum and vice versa, alternating at each point. The detailed proofs of each inequality are omitted here and can be found in Appendix B.2. Here, we present an overview of the construction. Fig. 1 gives an example for $m = 10, n = 5$ and Lorentzian peaks.

We denote the sums of the functions by f_m and f_n respectively. Both sums have one summand of significant width. This term of f_m has a width of w_1 and is centered around $c_1 = 0$. The corresponding summand of f_n is centered around a sufficiently large point $c_1 > 0$ with a width of $v_1 < w_1$. The two functions intersect twice. In Fig. 1, the two intersections lie at about $x = 297$ and $x = 309.5$ and can be seen in the bottom left plot.

The remaining intersections can be found within two intervals, namely in $[-m - w_1, -w_1]$ and in $[c_1 - v_1, c_1 + v_1]$. We define functions of very small widths so that the amount added to the sum outside a given range around their center is negligible. In the left interval, $n-1$ functions are defined in both sums, and each pair of functions contributes 4 intersections, while only marginally interfering with other peaks. This is illustrated in the top two plots of Fig. 1.

If $m > n$, we define the remaining $m-n$ functions of f_m in the right interval so that they each intersect the widest function of f_n twice. This can be seen in the bottom left of Fig. 1

Adding the first two intersections to the $4n-4$ found in the left interval and to the $2(m-n)$ found in the right interval leads to a total number of $2m+2n-2$ intersections.

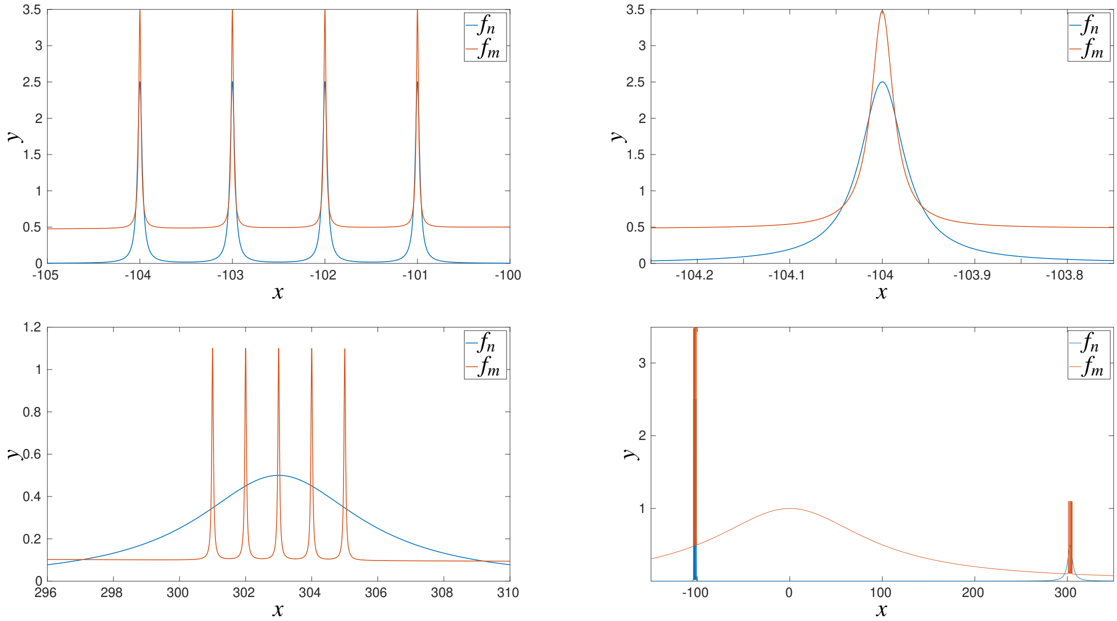

Figure 1: Two sums of Lorentzians with different parameters, showing the maximum number of intersection points.

3.1. Lorentz Functions

Again, we start with the simplest type of model function. Consider $m, n \in \mathbb{N}$, and a finite subset $X \subseteq \mathbb{R}$. Let $(c_i, v_i), (d_j, w_j)$ represent different parameter tuples and let $\lambda_i, \mu_j \in \mathbb{R}$ such that

$$\sum_{i=1}^{n} \frac{\lambda_i v_i^2}{v_i^2 + (x - c_i)^2} = \sum_{j=1}^{m} \frac{\mu_j w_j^2}{w_j^2 + (x - d_j)^2} \qquad \text{for } x \in X.$$

After multiplying both sides by both denominators the resulting polynomials have a degree of $2m + 2n - 2$ and the difference can have at most $2m+2n-2$ zeros. If more zeros existed, then the equality would hold for all $x \in \mathbb{R}$, which is impossible as per Thm. 1. This results in an upper bound of $k_L(m,n) \leq 2m + 2n - 2$.

To confirm if $f_{0,1} := L_{0,1}$ fulfills the requirements of Thm. 5, we can see that $L_{c,w}(x) = L_{0,1}\left(\frac{x-c}{w}\right)$. Additionally, conditions (I) - (III) and (V) are trivially satisfied. Assume $c \in \mathbb{R}$ and $w > 1$. Then, we have

$$\frac{L_{c,1}(x)}{L_{0,w}(x)} = \frac{w^2 + x^2}{w^2\left(1 + (x-c)^2\right)} \overset{x\to\infty}{\to} \frac{1}{w^2} < \frac{1}{w},$$

satisfying condition (IV) as well. By Thm. 5 we can deduce that $k_L(m,n) \geq 2m + 2n - 2$ and ultimately, we have $k_L(m,n) = 2m + 2n - 2$.

This proves that two sums of n and m Lorentz functions cannot be equal at $2m+2n-1$ or more points. Conversely, for a spectrum of 10000 points with 40 peaks present, a sum of 4960 or fewer Lorentz functions would be unable to replicate the spectrum at these points.

3.2. Gauss Functions

Again we would like to obtain a value for $k_G(m,n)$.

First, we must prove that the family of Gaussians, as defined in Eq. (2), satisfies the requirements of Thm. 5. Note that $G_{c,w}(x) = G_{0,1}\left(\frac{x-c}{w}\right)$. Also note that conditions (I) - (III) as well as (V) are trivially satisfied. If $c \in \mathbb{R}$ and $w > 1$, then

$$\frac{G_{c,1}(x)}{G_{0,w}(x)} = \exp\left(-\ln(2)\left(\left(1 - \frac{1}{w^2}\right)x^2 - 2cx + c^2\right)\right) \overset{x\to\infty}{\to} 0 < \frac{1}{w}.$$

Thus, Thm. 5 implies that $k_G(m,n) \geq 2m + 2n - 2$.

Second, for an upper bound, we refer to the following lemma.

Lemma 3.3. *For $m, n \in \mathbb{N}$ we have*

$$k_G(m,n) \leq 2^{m+n} - 2.$$

The proof of this lemma relies heavily on Rolle's theorem [17, App. A4]. According to this theorem, if a differentiable function $f : \mathbb{R} \to \mathbb{R}$ has at least n zeros $a_1 < \ldots < a_n$, then its derivative f' has at least $n-1$ zeros b_i with $a_i < b_i < a_{i+1}$. Conversely, if the derivative f' has no more than $n-1$ zeros, we can conclude that f must have at most n zeros. This argument can also be applied iteratively. Assuming that f is differentiable n-times and $f^{(n)}$ has no more than k zeros, then the number of zeros of f must be less than or equal to $k+n$. In the proof, we define a procedure for eliminating each summand by utilizing differentiation to obtain an upper bound to the number of zeros of a sum of scaled Gaussians. We also use the fact that for some polynomial $p(x)$ we have $\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)G_{c_i,v_i}(x)\right) = q(x)G_{c_i,v_i}(x)$, where $q(x)$ is another polynomial satisfying $\deg(q) \leq \deg(p) + 1$. The proof is omitted and can be found in Appendix B.3

This represents the only information on the upper limit of the number of intersections of sums of Gaussians. We can deduce that $2m + 2n - 2 \leq k_G(m,n) \leq 2^{m+n} - 2$. When $m = n = 1$, this implies that $k_G(1,1) = 2$, which can also be obtained by transforming $\lambda G_{c,v}(x) = \mu G_{d,w}(x)$ into a quadratic equation.

However, this result is of limited practical use. For a spectrum of 10000 points with 40 Gaussians present, we cannot exclude the possibility of a single Gaussian fitting the spectrum exactly at all 10000 points, since $2^{40+1} - 2 >$ 10000.

However, for a spectrum of 10000 points with 10 Gaussians present, we can exclude the possibility of 3 or fewer Gaussians fitting the spectrum due to $2^{10+3} - 2 = 8190 < 10000$.

3.3. Voigt Functions

It is currently unclear if $k_V(m,n) < \infty$. However, a lower bound can be found. Recall that

$$\begin{aligned} V_{c,v,w}(x) &= (G_{0,v} * L_{c,w})(x) \\ \lim_{v\searrow 0} V_{c,v,w}(x) &= \lim_{v\searrow 0}(G_{0,v} * L_{c,w})(x) = (\delta * L_{c,w})(x) = L_{c,w}(x), \end{aligned}$$

where δ is the Dirac delta distribution. This implies that the proof of Thm. 5 can also be conducted by using Voigt functions with sufficiently small parameters v. Therefore, we can achieve at least $2m+2n-2$ intersections using sums of m and n Voigt functions and $k_V(m,n) \geq 2m + 2n - 2$.

3.4. *Pseudo-Voigt Functions*

For all $m, n \in \mathbb{N}$ we can easily establish a lower bound to the number of intersections $k_{PV}(m, n) \geq 2m + 2n - 2$, by selecting $\lambda_i = \mu_j = 0$. The resulting functions are all Lorentzian and can have at least $2m + 2n - 2$ intersection points as proven in Thm. 5. Note that selecting $\lambda_i = \mu_j = 0$ results in a significantly simplified version of the function.

To find an upper bound, we refer to the following lemma.

Lemma 3.4. *For $m, n \in \mathbb{N}$, we have*

$$k_{PV}(m, n) \leq \left(2^{m+n+2} - 2\right)(m + n) - 2.$$

The proof can be found in Appendix B.4.

Numerical results suggest that neither the upper nor the lower bound is strict for any m, n. Even for $m = n = 1$, the bounds imply $2 \leq k_{PV}(1, 1) \leq 26$. Nonetheless, we can find parameters such that at least 6 intersections between two pseudo-Voigt functions exist and therefore $2 < 6 \leq k_{PV}(1, 1) \leq 26$, see Appendix C.4.

This result, again, is of limited use. For a spectrum of 10000 points consisting of 5 pseudo-Voigt peaks, we can only guarantee that three or fewer pseudo-Voigt functions cannot fit the spectrum at all points. For 4 functions, we have $(2^{5+4+2} - 2) \cdot (5 + 4) - 2 = 18412 > 10000$.

4. Conclusion

Modeling and simulation in natural sciences, engineering, medicine and other fields is an important tool for gaining a deeper understanding of the system behavior and for decision making. Model uniqueness is a valuable prerequisite for drawing the right conclusions from the simulated data. We hope that this work can contribute to a better understanding of the reliability and uniqueness of hard models for NMR spectral data and their qualitative and quantitative analysis. Such an analysis is a building block for successfully solving chemical structure elucidation problems.

In this work, we first state that the uniqueness of NMR hard models in terms of Lorentz, Gauss, pseudo-Voigt and Voigt functions with distinct center and width values requires mathematical analysis. Proofs are given for the linear independence of finite sets of these basis functions. Thus, any continuous NMR spectrum consisting of a finite sum of these functions can be uniquely hard modeled. For discrete spectra, we have introduced a minimum number of points $k_f(m, n) + 1$ required in order to prevent a sum of n model functions from being represented by a sum of m differently parametrized model functions and in most cases we provided upper and lower bounds on $k_f(m, n)$. Conversely, if the number of points of a spectrum consisting of n functions is given, we have a lower bound on the number m of model functions needed for an ambiguous hard model. However, it is uncertain whether an arbitrary set X allows ambiguous decompositions for $|X| \leq k_f(m, n)$, but it is known that there is some set X of this cardinality so that an ambiguous hard model exists. It is understood that the lower bound given by Thm. 5 is strict for Lorentz functions and not strict for pseudo-Voigt functions. However, it remains uncertain whether it is really strict for Gaussians. Numerical tests indicate strictness by randomly choosing parameters for both sums and then checking the number of intersections.

In summary, hard models of continuous NMR spectra are unique, and hard models of discrete NMR spectra can be assumed to be unique if the number of points at which the spectrum is recorded is high and if the number of basis functions used to build the model is not too large.

Appendix A. Two Fourier Transforms

In general, the Fourier transform of functions commonly used in signal theory is well known. In this appendix, we derive two formulas needed in this paper.

Lemma 2.2

Let $(c, w) \in \mathbb{R} \times \mathbb{R}_{>0}$ and $f = L_{c,w}$. By applying the Fourier transform to f and substituting $y = \frac{x-c}{w}$ we obtain

$$\begin{aligned}(\mathcal{F} f)(\omega) &= \int_{\mathbb{R}} \frac{w^2}{w^2 + (x-c)^2} \cdot \mathrm{e}^{-\mathbf{i}x\omega} \mathrm{d}x \\ &= w\mathrm{e}^{-\mathbf{i}c\omega} \cdot \int_{\mathbb{R}} \frac{1}{1+y^2} \cdot \mathrm{e}^{-\mathbf{i}y(w\omega)} \mathrm{d}y.\end{aligned}$$

The integral $\int_{\mathbb{R}} \frac{1}{1+y^2} \mathrm{e}^{-\mathbf{i}y\alpha} \mathrm{d}y = \pi \mathrm{e}^{-|\alpha|}$ is well known in complex analysis and is used to determine the characteristic function of a Cauchy distribution [10, App. D]. Consequently, we have

$$(\mathcal{F} f)(\omega) = \underbrace{w\pi}_{:=C(w)} \mathrm{e}^{-w|\omega|} \cdot \mathrm{e}^{-\mathbf{i}c\omega}.$$

□

Lemma 2.4

Let $(c, w) \in \mathbb{R} \times \mathbb{R}_{>0}$ and $f = G_{c,w}$. By applying the Fourier transform to f and substituting $y = \frac{\sqrt{\ln 2}}{w}(x-c)$ we obtain

$$\begin{aligned}(\mathcal{F} f)(\omega) &= \int_{\mathbb{R}} \mathrm{e}^{-\ln(2)\cdot\frac{(x-c)^2}{w^2}} \cdot \mathrm{e}^{-\mathbf{i}x\omega} \mathrm{d}x \\ &= \frac{w}{\sqrt{\ln 2}} \mathrm{e}^{-\mathbf{i}c\omega} \cdot \int_{\mathbb{R}} \mathrm{e}^{-y^2} \cdot \mathrm{e}^{-\mathbf{i}\frac{w}{\sqrt{\ln 2}} y\omega} \mathrm{d}y.\end{aligned}$$

To simplify the calculations, we write $\alpha := \frac{\mathbf{i}w\omega}{2\sqrt{\ln 2}}$. All that remains is to find a value for the integral

$$\int_{\mathbb{R}} \mathrm{e}^{-(y^2+2y\alpha)} \mathrm{d}y = \mathrm{e}^{\alpha^2} \int_{\mathbb{R}} \mathrm{e}^{-(y+\alpha)^2} \mathrm{d}y.$$

To calculate the integral on the right-hand side we define the function $g(\beta) = \int_{\mathbb{R}} \mathrm{e}^{-(y+\beta)^2} \mathrm{d}y : \mathbb{C} \to \mathbb{C}$. Now, for $\beta \in \mathbb{R}$, we have $g(\beta) = \sqrt{\pi}$ for this well-known integral. [11, Ex. 2.6]. Since g is an entire function obtained as an integral over another entire function, we have $g(\beta) = \sqrt{\pi}$ for all $\beta \in \mathbb{C}$ [14, Cor. 3.57].

This leads to the following equation

$$(\mathcal{F} f)(\omega) = \frac{w}{\sqrt{\ln 2}} \exp\left(\frac{-w^2\omega^2}{4\ln(2)} + -\mathbf{i}c\omega\right) \cdot \sqrt{\pi}.$$

Finally, we can define $w' = \frac{w}{2\sqrt{\ln(2)}} \in \mathbb{R}_{>0}$ and obtain

$$(\mathcal{F} f)(\omega) = \underbrace{\frac{\sqrt{\pi}w}{\sqrt{\ln 2}}}_{:=C(w)} \mathrm{e}^{-w'^2\omega^2} \cdot \mathrm{e}^{-\mathbf{i}c\omega}.$$

□

Appendix B. Proofs of mathematical theorems

This appendix contains those proofs that are omitted from the main body of the paper for better readability.

Appendix B.1. Proof of Lemma 2.3

We conduct this proof in two steps. First, we prove that there are some $\mu_j \in \mathbb{C}$ such that

$$\sum_{j=1}^{m} \mu_j e^{-\mathbf{i}c_j x} = 0 \tag{B.1}$$

holds. Second, we show that the functions $e^{-\mathbf{i}c_j x} : \mathbb{R} \to \mathbb{C}$ are linearly independent for different c_j. Therefore, all $\mu_j = 0$, which implies $\lambda_j = 0$ for all j.

We assume that $|c_1| = \max \left\{ |c_j| \,\middle|\, j \in [m] \right\}$. Furthermore, we have $c_1 \neq 0$, if $m \geq 2$.

First, we consider the case where $m = 1$ and $c_1 \neq 0$. We define the sequence $\omega_k = k \cdot \frac{2\pi}{c_1}$ for $k \in \mathbb{N}$, satisfying $e^{-\mathbf{i}c_1\omega_k} = 1$. From the convergence in (9) we can deduce that $\lambda_1 = 0$. In the case where $c_1 = 0$, (9) still implies $\lambda_1 = 0$.

Second, we consider the case where $m \geq 2$. For notational purposes, we define $f_j(\omega) = e^{-\mathbf{i}c_j\omega}$. Next, we examine the sequence $\left(\omega_k^{(1)}\right)_{k\in\mathbb{N}} := k \cdot \frac{2\pi}{c_1}$ satisfying

$$\lambda_1 f_1(\omega_k^{(1)}) = \lambda_1.$$

Additionally, we define $\mu_1 = \lambda_1$. The Bolzano-Weierstrass theorem implies the existence of a convergent subsequence of the bounded sequence $\lambda_2 f_2\left(\left(\omega_k^{(1)}\right)\right)$ converging to some $\mu_2 \in \mathbb{C}$. The corresponding subsequence of $\left(\omega_k^{(1)}\right)$ is denoted as $\left(\omega_k^{(2)}\right)$, which satisfies

$$\lambda_1 f_1\left(\omega_k^{(2)}\right) \overset{k\to\infty}{\to} \mu_1 \qquad \text{and} \qquad \lambda_2 f_2\left(\omega_k^{(2)}\right) \overset{k\to\infty}{\to} \mu_2.$$

If we define subsequences $\left(\omega_k^{(3)}\right), \ldots, \left(\omega_k^{(m)}\right)$ and limits $\mu_3, \ldots, \mu_m$ according to the above procedure, we can obtain the following convergences for all $j \in [m]$

$$\lambda_j f_j\left(\omega_k^{(m)}\right) \overset{k\to\infty}{\to} \mu_j.$$

Note that $0 \neq |\mu_j| = |\lambda_j|$, since the image of $\lambda_j f_j$ is a subset of $\{z \in \mathbb{C}, \mid |z| = |\lambda_j|\}$, which is a closed set. Additionally, let $x \in \mathbb{R}$. We define $\omega_k^{(m,x)} = \omega_k^{(m)} + x$ for all $k \in \mathbb{N}$ and note that

$$\lambda_j f_j\left(\omega_k^{(m,x)}\right) = \lambda_j f_j\left(\omega_k^{(m)}\right) e^{-\mathbf{i}c_j x} \overset{k\to\infty}{\to} \mu_j e^{-\mathbf{i}c_j x}.$$

By using (11) for all $x \in \mathbb{R}$, we obtain

$$\sum_{j=1}^{m} \lambda_j f_j\left(\omega_k^{(m,x)}\right) \overset{k\to\infty}{\to} \sum_{j=1}^{m} \mu_j e^{-\mathbf{i}c_j x} = 0.$$

To complete the proof, it suffices to show that the functions $e^{-\mathbf{i}c_j x}$ are linearly independent for different c_j. Let $x_1 := \frac{\pi}{2\cdot|c_1|} > 0$. This ensures that $x_1 c_j \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and thus, the values $e^{-\mathbf{i}c_j x_1}$ are distinct for different c_j. Additionally, let

$$x_2 = 2x_1, \quad x_3 = 3x_1, \quad \ldots, \quad x_{m-1} = (m-1)x_1, \quad x_m = 0.$$

Evaluating Eq. (B.1) at the points $x_1, \ldots, x_m$ yields a homogeneous linear system of equations with the following matrix

$$V = \begin{pmatrix} e^{-\mathbf{i}c_1 x_m} & \cdots & e^{-\mathbf{i}c_m x_m} \\ e^{-\mathbf{i}c_1 x_1} & \cdots & e^{-\mathbf{i}c_m x_1} \\ e^{-\mathbf{i}c_1 x_2} & \cdots & e^{-\mathbf{i}c_m x_2} \\ \vdots & \ddots & \vdots \\ e^{-\mathbf{i}c_1 x_{m-1}} & \cdots & e^{-\mathbf{i}c_m x_{m-1}} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ e^{-\mathbf{i}c_1 x_1} & \cdots & e^{-\mathbf{i}c_m x_1} \\ \left(e^{-\mathbf{i}c_1 x_1}\right)^2 & \cdots & \left(e^{-\mathbf{i}c_m x_1}\right)^2 \\ \vdots & \ddots & \vdots \\ \left(e^{-\mathbf{i}c_1 x_1}\right)^{m-1} & \cdots & \left(e^{-\mathbf{i}c_m x_1}\right)^{m-1} \end{pmatrix}.$$

This Vandermonde matrix has $\det(V) = \prod_{j<k}\left(e^{-\mathbf{i}c_k x_1} - e^{-\mathbf{i}c_j x_1}\right)$, [5, 4.6]. Since the values $e^{\mathbf{i}c_j x_1}$ are distinct, we have $\det(V) \neq 0$. Therefore, all $\mu_j = 0$. The fact $|\mu_j| = |\lambda_j|$ implies the desired result. □

Appendix B.2. Proof of Theorem 5

From conditions (I) - (V) we obtain the following auxiliary statements:

A: We have $f_{0,1}(x) \overset{x\to\infty}{\to} 0$, using $f_{0,1} \in L_1(\mathbb{R})$ and condition (III).

B: For every $w \in \mathbb{R}_{>0}$ there exists some $x^* \in \mathbb{R}_{>0}$ such that $f_{0,w}(x^*) < \frac{1}{10}$. This is because statement A implies $f_{0,w}(x) = f_{0,1}\left(\frac{x}{w}\right) \overset{x\to\infty}{\to} 0$.

C: There exists some $\delta \in \mathbb{R}_{>0}$ such that $f_{0,\delta}\left(\frac{1}{2}\right) < \frac{1}{30m}$. This is true due to the fact that statement A implies $f_{0,\delta}\left(\frac{1}{2}\right) = f_{0,1}\left(\frac{1}{2\delta}\right) \overset{\delta\to 0}{\to} 0$.

D: Assume $v \leq w$, conditions (II) and (III) imply the relation $f_{0,v}(x) \leq f_{0,w}(x)$ for all x.

E: For every $m \in \mathbb{N}$ there exists $w' \in \mathbb{R}_{>0}$ such that $f_{0,w'}(x) \in [0.4, 0.5]$ for $x \in [w', w' + m]$. Based on conditions (I), (III), (V) and statement A there exists a point $y > 1$ where $f_{0,1}(y) = 0.4$. We define $w' = \frac{m}{y-1}$ and obtain

$$f_{0,w'}(w' + m) = f_{0,1}\left(1 + \frac{m}{w'}\right) = 0.4.$$

Note that for $w \geq w'$ statement E also holds.

F: For all $c \in \mathbb{R}$, $w \in \mathbb{R}_{>0}$ and $x \in [c - w, c + w]$ we have $f_{c,w}(x) \geq \frac{1}{2}$ based on conditions (I) - (III) and (V).

G: For each $\delta > 0$ there exists some $N \in \mathbb{N}$ such that $f_{0,\delta/N}(\delta) < \frac{1}{10}$. This holds because statement A implies $f_{0,\delta/N}(\delta) = f_{0,1}(N) \overset{N\to\infty}{\to} 0$.

We frequently refer to these statements and denote them by their abbreviated definitions $A - G$.

Let $m \geq n$ be natural numbers. For $i \in [n]$ and for $j \in [m]$ let $f_{c_i,v_i}(x)$ and $f_{d_j,w_j}(x)$ be functions parametrized as defined in Thm. 5 and let $f_{0,1}$ satisfy conditions (I) to (V). We need to define the parameters c_i, d_j, v_i and w_j, and prove that the sums

$$f_n(x) = \sum_{i=1}^{n} \lambda_i f_{c_i,v_i}(x) \quad \text{and} \quad f_m(x) = \sum_{j=1}^{m} \mu_j f_{d_j,w_j}(x) \tag{B.2}$$

are equal in at least $2m + 2n - 2$ different points $x \in X$.

Let w' be sufficiently large such that $f_{0,w'}(x) \in [0.4, 0.5]$ for $x \in [w', w' + m]$ and let $\rho = 3mn$. We define $w_1 = \max\{m, w', \rho + 1\}$ and choose x^* sufficiently large such that $f_{0,w_1}(x^*) < \frac{1}{10}$. Let δ be sufficiently small such that $f_{0,\delta}\left(\frac{1}{2}\right) < \frac{1}{30m}$. Note that $\delta < \frac{1}{2}$. Then, let N be sufficiently large such that $f_{0,\delta/N}(\delta) < \frac{1}{2}$. The previous definitions use statements E, B, C and G respectively. Additionally, we define the following parameters

$c_1 = x^* + \frac{1}{2}(m - n + 1),$	$v_1 = \frac{1}{2}(m - n + 1),$	$\lambda_1 = \frac{1}{2}$	
$c_i = -w_1 - i + 1,$	$v_i = \delta,$	$\lambda_i = \frac{5}{2}$	for $i = 2, \ldots, n$
$d_1 = 0,$	w_1 prev. defined	$\mu_1 = 1$	
$d_j = -w_1 - j + 1,$	$w_j = \frac{\delta}{N},$	$\mu_j = 2$	for $j = 2, \ldots, n$
$d_j = x^* + j - n,$	$w_j = \delta,$	$\mu_j = 1$	for $j = n + 1, \ldots, m.$

Note that these parameters yield functions that mainly intersect in two intervals. Firstly, in $[-w_1 - m, -w_1]$ and secondly, in $[x^*, x^* + m - n + 1]$. Additionally, there are two more intersection points. One is within the interval $(-w_1 - \frac{1}{2}, x^* + \frac{1}{2})$ and the other is greater than $x^* + m - n + 1$.

In total, we define $2m + 2n - 1$ points at which $f_n < f_m$ and $f_m < f_n$ alternate. Since both functions are continuous, according to the intermediate value theorem [13] there must be at least $2m + 2n - 2$ points $x \in X$, where $f_n(x) = f_m(x)$.

Part 1: $[-w_1 - m, -w_1]$
The first $4n-3$ points are defined as follows

$$x_0 = -w_1 - n + \frac{1}{2}, \qquad x_4 = -w_1 - n + \frac{3}{2}, \qquad \dots, \qquad x_{4n-4} = -w_1 - \frac{1}{2} \tag{B.3}$$
$$x_1 = -w_1 - n + 1 - \delta, \qquad x_5 = -w_1 - n + 2 - \delta, \qquad \dots, \qquad x_{4n-7} = -w_1 - 1 - \delta \tag{B.4}$$
$$x_2 = -w_1 - n + 1, \qquad x_6 = -w_1 - n + 2, \qquad \dots, \qquad x_{4n-6} = -w_1 - 1 \tag{B.5}$$
$$x_3 = -w_1 - n - 1 + \delta, \qquad x_7 = -w_1 - n - 2 + \delta, \qquad \dots, \qquad x_{4n-5} = -w_1 - 1 + \delta. \tag{B.6}$$

All points lie within the given interval and based on (II), (V), D and E we have $f_{0,w_1}(x_k) \in [0.4, 0.5]$. Let x be a point defined by Eqs. (B.3). Note that $v_1 < m \leq w_1$ and $c_1 \geq 0$, resulting in

$$\begin{aligned} f_{c_1,v_1}(x) &\overset{(III)}{\leq} f_{c_1,v_1}(0) \overset{(II)}{=} f_{0,v_1}(c_1) \\ &\overset{D}{\leq} f_{0,w_1}\left(x^* + \frac{1}{2}(m-n+1)\right) \overset{(III)}{\leq} f_{0,w_1}(x^*) \overset{B}{<} \frac{1}{10}. \end{aligned}$$

We observe that $|x - c_i| \geq \frac{1}{2}$ for all $i = 2, \dots, n$ and obtain

$$\begin{aligned} f_n(x) &= \frac{1}{2} \cdot f_{c_1,v_1}(x) + \sum_{i=2}^{n} \frac{5}{2} \cdot f_{c_i,v_i}(x) \\ &\overset{B,C}{<} \frac{1}{20} + \frac{5(n-1)}{2 \cdot 30m} < \frac{2}{15} < \frac{2}{5} \leq f_{0,w_1}(x) \leq f_m(x). \end{aligned}$$

Note that all functions are nonnegative.

Let x be a point defined in Eqs. (B.4), where $x = -10m - k - \delta$ for some $k = 1, \dots, n-1$. Additionally, let $c_l = d_l = -10m - k$ for some $l \in \{2, \dots, n\}$. It follows that $\frac{5}{2} f_{c_l,\delta}(c_l - \delta) = \frac{5}{2} f_{0,1}(-1) = \frac{5}{4}$ and $f_n(x) \geq \frac{5}{4}$. Similarly, we have

$$3 f_{d_l,\delta/N}(c_l - \delta) = 3 f_{0,\delta/N}(-\delta) \overset{G}{<} \frac{3}{10}.$$

Moreover, we again have $|x - d_j| \geq \frac{1}{2}$ for $j \in [m] \setminus \{1, l\}$. Finally, we have

$$\begin{aligned} f_m(x) &= 1 \cdot f_{d_1,w_1}(x) + 3 \cdot f_{d_l,w_l}(x) + \sum_{j \in [m] \setminus \{1,l\}} \mu_j \cdot f_{d_j,w_j}(x) \\ &\overset{E,G,C}{\leq} \frac{1}{2} + \frac{3}{10} + \frac{3(m-2)}{30m} < \frac{9}{10} < \frac{5}{4} \leq f_n(x). \end{aligned}$$

The above inequality holds for all x defined by Eqs. (B.6) based on the same arguments. It is important to note that $f_{c_l,\delta}(c_l - \delta) = f_{c_l,\delta}(c_l + \delta)$.

Finally, suppose x is a point defined in Eqs. (B.5), where $x = -10m - k$ for some $k = 1, \dots, n-1$ and let $c_l = d_l = -10m - k$. Following similar arguments, we obtain

$$\begin{aligned} f_n(x) &= \frac{1}{2} f_{c_1,v_1}(x) + \frac{5}{2} f_{c_l,v_l}(x) + \sum_{i \in [n] \setminus \{1,l\}} \lambda_i f_{c_i,v_i}(x) \\ &\overset{B,C}{\leq} \frac{1}{20} + \frac{5}{2} + \frac{5(n-2)}{2 \cdot 30m} < \frac{158}{60} < 3 = 3 \cdot f_{d_l,w_l}(x) \leq f_m(x). \end{aligned}$$

Part 2: $[x^*, x^* + m - n + 1]$
To begin the second part of the proof, we define

$$y_0 = x^* + \frac{1}{2}, \qquad y_2 = x^* + \frac{3}{2}, \qquad \dots, \qquad y_{2m-2n} = x^* + m - n - \frac{1}{2} \tag{B.7}$$
$$y_1 = x^* + 1, \qquad y_3 = x^* + 2, \qquad \dots, \qquad y_{2m-2n-1} = x^* + m - n - 1. \tag{B.8}$$

Note that all $y_k \in [c_1 - v_1, c_1 + v_1]$.

We use the same procedure as in the first part of the proof. Let y be a point as defined by Eqs. (B.7). By design, the distance from y to any peak center except for c_1 is at least $\frac{1}{2}$. We can conclude that

$$\begin{aligned} f_m(y) &= f_{d_1,w_1}(y) + \sum_{j=2}^{n} \mu_j f_{d_j,w_j}(y) \\ &\overset{B,C}{\leq} \frac{1}{10} + \frac{1}{10} < \frac{1}{4} \overset{F}{\leq} \frac{1}{2} f_{c_1,v_1}(y) \leq f_n(y). \end{aligned}$$

Let y be a point as defined by Eqs. (B.8). Again, there exists an l such that $y = d_l$, and the distance from y to the centers of other peaks is at least $\frac{1}{2}$. This implies that

$$\begin{aligned} f_n(y) &= \frac{1}{2} f_{c_1,v_1}(y) + \sum_{i=2}^{n} \frac{5}{2} f_{c_i,v_i}(y) \\ &\overset{C}{<} \frac{1}{2} + \frac{1}{12} < 1 = f_{d_l,w_l}(y) \leq f_m(y). \end{aligned}$$

Note that $f_n(x_{4n-4}) < f_m(x_{4n-4})$ and $f_m(y_0) < f_n(y_0)$.

Part 3: $[x^* + m - n + 1, \infty)$

To conclude the proof, our goal is to demonstrate the existence of some $x_r \in \mathbb{R}$ where $f_n(x_r) < f_m(x_r)$. Utilizing condition (IV), we have

$$\begin{aligned} \lim_{x\to\infty} \frac{f_n}{f_{0,w_1}}(x) &= \lim_{x\to\infty} \sum_{i=1}^{n} \frac{\lambda_i f_{c_i,v_i}(x)}{f_{0,w_1}(x)} = \lim_{x\to\infty} \sum_{i=1}^{n} \frac{\lambda_i f_{c_i,1}(x \cdot v_i)}{f_{0,w_1/v_i}(x \cdot v_i)} \\ &\overset{(IV)}{=} \sum_{i=1}^{n} \lambda_i \cdot \vartheta_{w_1/v_i} < \sum_{i=1}^{n} \frac{\lambda_i v_i}{w_1} < \frac{3mn}{w_1}. \end{aligned}$$

The preceding arguments show that as x approaches infinity, the function $\frac{f_n}{f_{d_1,w_1}}(x)$ approaches a value $\frac{\rho}{w_1} < 1$. Consequently, there exists some $x_r > x^* + m - n + 1$ at which the following inequality holds true

$$\frac{f_n}{f_{d_1,w_1}}(x_r) < 1 \leq \frac{f_m}{f_{d_1,w_1}}(x_r).$$

Conditions (1) and (3), along with statement A, imply that $f_{c,w}(x) \geq 0$. Therefore, $f_n(x_r) < f_m(x_r)$.

Ultimately, we have identified a set of $(4n-3) + (2m-2n+1) + 1 = 2m + 2n - 1$ points, namely

$$x_0 < \ldots < x_{4n-4} < y_0 < \ldots < y_{2m-2n} < x_r$$

at which $f_n < f_m$ and $f_m > f_n$ alternate, implying there are at least $2m + 2n - 2$ points $x \in X$ with $f_n(x) = f_m(x)$.

Appendix B.3. Proof of Lemma 3.3

Let $m, n \in \mathbb{N}$ and let $p_i = (c_i, v_i)$ and $q_j = (d_j, w_j)$ be parameter tuples. Further, let $\lambda_i, \mu_j \in \mathbb{R}$ for $i \in [n], j \in [m]$. Note that assuming positive coefficients λ_i, μ_j does not improve the upper bound in this proof. Therefore, we define $\lambda_{n+j} = -\mu_j$ and $p_{n+j} = q_j$. For some $x \in X$ we consider the following equation

$$\sum_{i=1}^{n} \lambda_i G_{p_i}(x) = \sum_{j=1}^{m} \mu_j G_{q_j}(x)$$

$$\sum_{i=1}^{m+n} \lambda_i G_{p_i}(x) = \sum_{i=1}^{m+n} p_{0,i}^{(0)}(x) \mathrm{e}^{q_i^{(0)}(x)} = 0.$$

In the above equation, $q_i^{(0)}(x)$ denotes the corresponding quadratic polynomial and $p_{d,i}^{(0)}$ denotes the corresponding polynomial of degree d. The upper index $p^{(j)}$ denotes the iteration step.

If we multiply both sides of the equation by $\mathrm{e}^{-q_1^{(0)}(x)}$ we obtain

$$\sum_{i=2}^{m+n} p_{0,i}^{(0)}(x)\mathrm{e}^{q_i^{(1)}(x)} + p_{0,1}^{(0)}(x) = 0.$$

Let us define the left-hand side as a differentiable function $f(x)$ and examine the zeros of its derivative $f'(x)$

$$f'(x) = \sum_{i=2}^{m+n} p_{1,i}^{(1)}(x)\mathrm{e}^{q_i^{(1)}(x)} = 0.$$

Note that we now have one less summand. We proceed by multiplying by the inverse of the first remaining exponential term $\mathrm{e}^{-q_2^{(1)}(x)}$ resulting in

$$\sum_{i=3}^{m+n} p_{1,i}^{(1)}\mathrm{e}^{q_i^{(2)}(x)} + p_{1,2}^{(1)} = 0.$$

The function $g(x)$ on the left-hand side shares its zeros with $f'(x)$ and is also smooth. Thus, we differentiate twice and examine the number of solutions of

$$g''(x) = \sum_{i=3}^{m+n} p_{3,i}^{(2)}(x)\mathrm{e}^{q_i^{(2)}(x)} = 0.$$

After each iteration step, the number of summands decreases by one, allowing us to repeat the procedure. Upon iterating the process k times, we obtain

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{2^{k-1}} r(x) = \sum_{i=k+1}^{m+n} p_{(2^k-1,i)}^{(k)}(x)\mathrm{e}^{q_i^{(k)}(x)} = 0.$$

To eliminate the next term, we can multiply the above equation by $\mathrm{e}^{-q_{k+1}^{(k)}(x)}$ and observe that the resulting function has the same set of zeros as the previous one. Differentiating 2^k more times lets the polynomial $p_{(2^k-1,k+1)}^{(k)}(x)$ vanish and the resulting polynomial factors of the other terms have a degree of $2^k - 1 + 2^k = 2^{k+1} - 1$.

After $m + n - 1$ steps we ultimately arrive at

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{2^{m+n-2}} z(x) = p_{(2^{m+n-1}-1,m+n)}^{(m+n-1)}(x)\mathrm{e}^{q_{m+n}^{(m+n-1)}(x)} = 0$$

This equation has at most $2^{m+n-1} - 1$ solutions, indicating that $z(x)$ has at most $2^{m+n-1} - 1 + 2^{m+n-2}$ zeros by Rolle's theorem. Iterating backwards yields that the original function f has at most

$$\left(\sum_{i=1}^{m+n-1} 2^i\right) - 1 = 2^{m+n} - 2$$

zeros.

Appendix B.4. Proof of Lemma 3.4

We conduct this proof similarly to the proof of Lemma 3.3 but omit some details.

For $i \in [n]$ and $j \in [m]$ let $PV_{p_i}(x)$ and $PV_{q_j}(x)$ be pseudo-Voigt functions and let $\alpha_i, \beta_j \in \mathbb{R}$, with parameters $p_i = (c_i, v_i, \lambda_i), q_j = (d_j, w_j, \mu_j)$. Assuming non-negative α_i, β_j currently does not improve the upper bound. Therefore we define $\alpha_{n+j} = -\beta_j$ and $p_{n+j} = q_j$. For some $x \in X$ we consider the equation

$$\sum_{i=1}^{m+n} \alpha_i\lambda_i\mathrm{e}^{-\ln(2)\frac{(x-c_i)^2}{v_i}} + \alpha(1-\lambda_i)\frac{v_i^2}{v_i^2 + (x-c_i)^2} = 0.$$

Multiplying both sides by $\prod_{i=1}^{n+m} v_i^2 + (x - c_i)^2$ results in

$$\left(\sum_{i=1}^{m+n} p^{(0)}_{(2m+2n,i)}(x) \mathrm{e}^{q_i^{(1)}(x)} \right) + r(x) = 0$$

with $p^{(j)}_{d,i}$ being the corresponding polynomial of the degree d at step j of the iteration, $q_i^{(j)}$ being the corresponding quadratic polynomial at step j and $r(x)$ being a polynomial of degree $2m + 2n - 2$. We may define the left-hand side as a function $f(x)$ and examine the zeros of its derivative

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} \right)^{2m+2n-1} f(x) = \sum_{i=1}^{m+n} p^{(1)}_{(4m+4n-1,i)}(x) \mathrm{e}^{q_i^{(1)}(x)} = 0.$$

This is the initial step of our iteration. In step j, we multiply the equation by $\mathrm{e}^{-q_k^{(j)}(x)}$ and differentiate the resulting function $((m + n) \cdot 2^{k+1})$ times, akin to the proof of Lemma 3.3.

Upon completing $m + n - 1$ steps we arrive at

$$p^{(m+n)}_{((m+n)2^{m+n+1}-1,m+n)}(x) \mathrm{e}^{q^{(m+n)}_{m+n}(x)} = 0.$$

This equation has at most $(m + n) \cdot 2^{m+n+1} - 1$ zeros. By utilizing Rolle's theorem, we can determine an upper limit for the number of zeros of the initial equation by adding up the number of differentiations performed throughout the process. The total number of differentiations performed is

$$\begin{aligned} \sum_{k=1}^{m+n-1} \left[(m + n)2^{k+1} \right] + 2m + 2n - 1 &= \left(\sum_{k=0}^{m+n-1} 2^k \right) 2(m + n) - 1 \\ &= (2^{m+n+1} - 2)(m + n) - 1. \end{aligned}$$

Adding the two numbers yields a maximum number of

$$\left(2^{m+n+2} - 2 \right) (m + n) - 2$$

intersections between m and n pseudo-Voigt functions.

Appendix C. Illustrations and minor proofs

In this appendix we want to give some illustrations of the notions of practical and minimal ambiguity. In addition, we constructively prove that $k_{PV}(1, 1) \geq 6$ and prove the small remark made when defining Voigt functions in Eq. (3).

Appendix C.1. Practical and Theoretical Ambiguity

For distinct parameter tuples (c_i, w_i), Cor. 2.8 implies that the function $s(x)$ can be uniquely decomposed by

$$s(x) = \sum_{i=1}^{n} \alpha_i PV_{c_i,w_i,\lambda_i}(x)$$

in an exact, analytical sense. If we were to define the *practical ambiguity* of spectral hard models, it would be a case where there is some $\varepsilon > 0$ for which the following inequality holds for all $x \in [a, b]$

$$\left| \sum_{i=1}^{n} \alpha_i PV_{c_i,v_i,\lambda_i}(x) - \sum_{j=1}^{m} \beta_j PV_{d_j,w_j,\mu_j}(x) \right| < \varepsilon.$$

In this setting there is no practical uniqueness, even if ε is chosen arbitrarily small, due to the continuity of pseudo-Voigt (and other) model functions with respect to their parameters. Thus, different parameters can always be chosen that are very close to the original ones.

Furthermore, Fig. C.2 demonstrates practical ambiguities with significantly different parameters. Since the maximum deviation at any point is 0.0026, assuming there is a noise level of 1%, both solutions would be indistinguishable in these plots and practically ambiguous with $\varepsilon = 0.01$.

Appendix C.2. Construction of Minimal Ambiguous Spectra

We again consider Fig. C.2. Both fits of the spectrum have identified the leftmost peak almost identically. Therefore, we can construct another spectrum with an ambiguous hard model and fewer model functions, as shown in Fig. C.3.

Appendix C.3. A Different Representation of Voigt Functions

We demonstrate that the Voigt function resulting from $V_{c,d,v,w}(x) = G_{c,v} * L_{d,w}$ where $c, d \in \mathbb{R}$, and $v, w \in \mathbb{R}_{>0}$ can be represented using the format defined in (3). First, for $c', d' \in \mathbb{R}$ we have

$$G_{c,v}(x + c') = \exp\left(-\ln(2) \cdot \frac{(x + c' - c)^2}{v^2}\right) = G_{c-c',v}(x)$$
$$L_{d,w}(x + d') = \frac{w^2}{w^2 + (x + d' - d)^2} = L_{d-d',w}(x).$$

Second, we get

$$\begin{aligned}(G_{c,v} * L_{d,w})(x) &= \int_{\mathbb{R}} G_{c,v}(y)L_{d,w}(x - y)\mathrm{d}y \\ &= \int_{\mathbb{R}} G_{c,v}(y + c)L_{d,w}(x - y - c)\mathrm{d}y \\ &= \int_{\mathbb{R}} G_{0,v}(y)L_{c+d,w}(x - y)\mathrm{d}y \\ &= (G_{0,v} * L_{c+d,w})(x).\end{aligned}$$

This proves that including an additional center parameter for the Gaussian does not generalize the definition of the Voigt function. □

Appendix C.4. Proof that $k_{PV}(1, 1) \geq 6$

Let us define the two functions $f(x) = 1 \cdot PV_{0,1,1/2}(x)$ and $g(x) = \frac{13}{25} \cdot PV_{0,2,3/4}(x)$. Next, we evaluate these functions at the following points.

$\underline{x_1 = 0:}$

We have $f(0) = 1 > g(0) = \frac{13}{25}$.

$\underline{x_2 = 2:}$

We have

$$\begin{aligned}f(2) &= \frac{1}{32} + \frac{1}{10} < \frac{1}{5} \\ &< \frac{13}{50} = \frac{39}{200} + \frac{13}{200} = g(2).\end{aligned}$$

$\underline{x_3 = 8:}$

We have

$$\begin{aligned}170 &> \frac{1}{2} + 169 = \frac{2^2 \cdot (2^4)^2 \cdot 2^5}{2^{16}} + 169 > \frac{3 \cdot 13^2 \cdot 17}{2^{16}} + 169 \\ \frac{1}{130} &> \frac{39}{100} \cdot 2^{-16} + \frac{13}{1700} \\ f(8) = 2^{-65} + \frac{1}{130} &> \frac{39}{100} \cdot 2^{-16} + \frac{13}{100} \cdot \frac{1}{17} = g(8).\end{aligned}$$

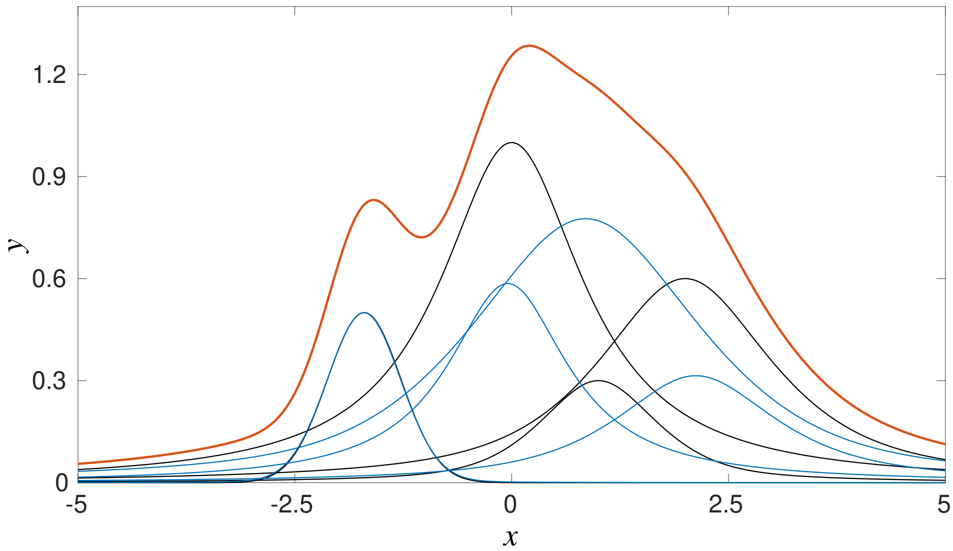

Figure C.2: Two groups of 4 pseudo-Voigt functions (black and blue) with distinct parameters. The sums of the two groups (red) have a maximum deviation of 0.0026 and an average deviation of 0.00072. Note that the leftmost peak is nearly identical in both sets.

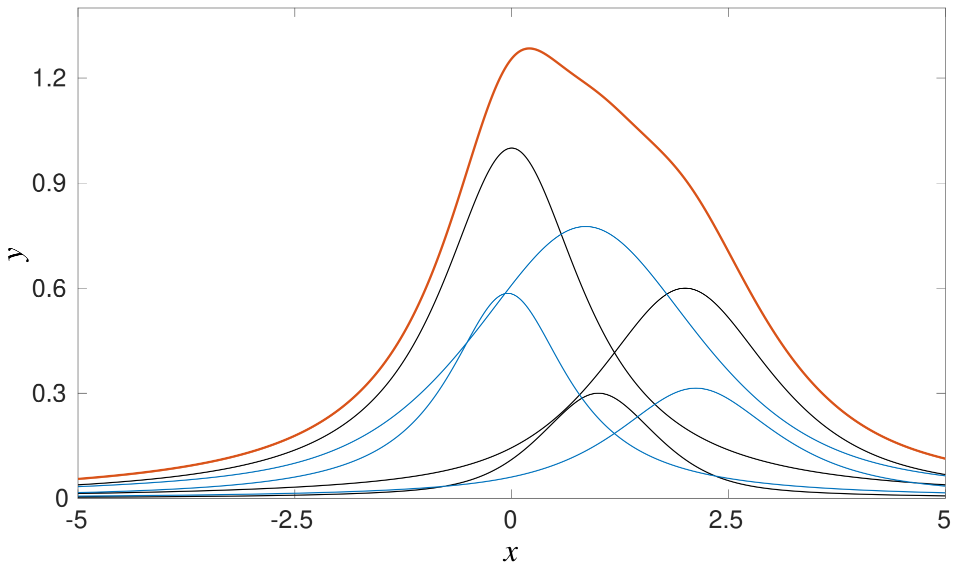

Figure C.3: Two groups of 3 pseudo-Voigt functions (black and blue) with distinct parameters. This spectrum is derived from Fig. C.2 by subtracting the leftmost peak. The sums of the two groups (red) now have a slightly larger maximum deviation of 0.0039 and an average deviation of 0.0015.

Finally, we consider the asymptotic behaviour where

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{\frac{1}{2}\cdot \exp\left(-\ln(2)x^2\right) + \frac{1}{2}\frac{1}{1+x^2}}{\frac{39}{100}\exp\left(-\ln(2)\frac{x^2}{4}\right) + \frac{13}{100}\frac{4}{4+x^2}}$$

$$= \lim_{x\to\infty} \frac{\frac{1}{2(1+x^2)}}{\frac{52}{100(4+x^2)}} = \lim_{x\to\infty} \frac{100(4+x^2)}{104(1+x^2)} = \frac{25}{26} < 1.$$

From this limit, we can conclude that there exists some $x_4 > x_3$ such that $f(x_4) < g(x_4)$. Using the symmetry of both functions, we know that the sequence $f(-x_4) < g(-x_4), f(-x_3) > g(-x_3), \ldots, f(x_4) < g(x_4)$ alternates. Both functions are continuous and according to the intermediate value theorem, there must be at least six points where $f(x) = g(x)$. Therefore, $k_{PV}(1,1) \geq 6$. □

References

[1] S.M. Abrarov and B.M. Quine. A rational approximation for efficient computation of the Voigt function in quantitative spectroscopy. *J. Math. Res.*, 7(2), 2015.

[2] F. Alsmeyer, H.-J. Koß, and W. Marquardt. Indirect spectral hard modeling for the analysis of reactive and interacting mixtures. *Appl. Spectrosc.*, 58(8):975–985, 2004.

[3] O. Christensen and K.L. Christensen. Linear independence and series expansions in function spaces. *Am. Math. Mon.*, 113(7):611–627, 2006.

[4] I. Daubechies. *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[5] G.H. Golub and C.F. Van Loan. *Matrix computations*. John Hopkins University Press, Baltimore, 1996.

[6] H.-W. Koh. *Feature extraction in NMR data analysis*. PhD thesis, TU Dortmund, Fachbereich Informatik, 2010.

[7] D.-W. Li, A.L. Hansen, C. Yuan, L. Bruschweiler-Li, and R. Brüschweiler. DEEP picker is a deep neural network for accurate deconvolution of complex two-dimensional NMR spectra. *Nat. Commun.*, 12(1):5229, 2021.

[8] W.F. Maddams. The scope and limitations of curve fitting. *Appl. Spectrosc.*, 34(3):245–267, 1980.

[9] D. Meinhardt, H. Schröder, J. Hellwig, E. Steimers, A. Friebel, T. Beweries, M. Sawall, E. von Harbou, and K. Neymeyr. Model-based signal tracking in the quantitative analysis of time series of NMR spectra. *J. Magn. Reson.*, 339:107212, 2022.

[10] W.R. Pestman. *Mathematical statistics*. Walter de Gruyter, Berlin, 2009.

[11] G. Plonka, D. Potts, G. Steidl, and M. Tasche. *Numerical fourier analysis*. Birkhäuser, Basel, 2018.

[12] D.A. Ramsay. Intensities and shapes of infrared absorption bands of substances in the liquid phase. *J. Am. Chem. Soc.*, 74(1):72–80, 1952.

[13] S.B. Russ. A translation of Bolzano's paper on the intermediate value theorem. *Hist. Math.*, 7(2):156–185, 1980.

[14] D.A. Salamon. *Funktionentheorie*. Springer-Verlag, Berlin, 2011.

[15] M. Schmid, H.-P. Steinrück, and J.M. Gottfried. A new asymmetric Pseudo-Voigt function for more efficient fitting of XPS lines. *Surf. Interface Anal.*, 46(8):505–511, 2014.

[16] N. Schmid, S. Bruderer, F. Paruzzo, G. Fischetti, G. Toscano, D. Graf, M. Fey, A. Henrici, V. Ziebart, B. Heitmann, et al. Deconvolution of 1D NMR spectra: a deep learning-based approach. *J. Magn. Reson.*, 347:107357, 2023.

[17] G.F. Simmons. *Calculus with analytic geometry*. McGraw Hill Education, New York, 1996.