



## EASY-GOING deconvolution: Combining accurate simulation and evolutionary algorithms for fast deconvolution of solid-state quadrupolar NMR spectra

Dennis L.A.G. Grimminck\*, Ben J.W. Polman, Arno P.M. Kentgens, W. Leo Meerts

Radboud University, Institute for Molecules and Materials, Heyendaalseweg 135, NL-6525 AJ Nijmegen, The Netherlands

### ARTICLE INFO

#### Article history:

Received 23 March 2011

Revised 21 April 2011

Available online 4 May 2011

#### Keywords:

Spectrum deconvolution

Solid-state NMR

Evolutionary algorithms

Quadrupolar NMR

### ABSTRACT

A fast and accurate fit program is presented for deconvolution of one-dimensional solid-state quadrupolar NMR spectra of powdered materials. Computational costs of the synthesis of theoretical spectra are reduced by the use of libraries containing simulated time/frequency domain data. These libraries are calculated once and with the use of second-party simulation software readily available in the NMR community, to ensure a maximum flexibility and accuracy with respect to experimental conditions. *EASY-GOING deconvolution (EGdeconv)* is equipped with evolutionary algorithms that provide robust many-parameter fitting and offers efficient parallelised computing. The program supports quantification of relative chemical site abundances and (dis)order in the solid-state by incorporation of (extended) Czek and order parameter models. To illustrate *EGdeconv*'s current capabilities, we provide three case studies. Given the program's simple concept it allows a straightforward extension to include other NMR interactions. The program is available as is for 64-bit Linux operating systems.

© 2011 Elsevier Inc. All rights reserved.

### 1. Introduction

Quadrupolar nuclear magnetic resonance is a valuable tool in materials research. The technique applies to atomic nuclei that possess a total angular momentum  $\geq 1$  and thereby a quadrupolar moment [1]. This concerns over two-thirds of the NMR active nuclei, hence the area of application for quadrupolar NMR is potentially large. Solid-state NMR techniques help to study the interaction between the quadrupolar moment of nuclear charge and the electrical field gradient (*efg*) generated by the chemical surroundings; the microscopic system of a nucleus and its surroundings is known as a site. The nucleus thereby acts as a probe for local structure, which typically extends over several coordination spheres [2]. This makes quadrupolar NMR a particularly sensitive technique for structural information from (partly) disordered solid-state materials, e.g. glasses, semiconductors and polymers, for example materials that cannot be studied by X-ray techniques due to their lack of long range order.

Structural information contained in the experimental spectra can be extracted by deconvolution in terms of quadrupolar interaction parameters and relative abundances of different sites. This requires a deconvolution model, which accurately describes the experiment and spin system at reasonable computational costs. Fitting programs such as DMFit [3] and QuadFit [4] have their own dedicated model, which is optimised towards low costs, but not versatility in terms of experimental conditions. Conversely,

there are general purpose NMR simulation programs such as SIMPSON [5] or SPINEVOLUTION [6], that are versatile in their capabilities to mimic experimental conditions and varying spin systems, but miss dedication for spectrum fitting. In practise the quality of simulations of quadrupolar NMR spectra is mostly restrained by computational costs, for example due to the need for orientational averaging of crystallites for powder spectra. With this in mind we created a separation between the computationally intensive simulations and the actual data-fitting, by the use of second-party software, in the work presented here SIMPSON, to pre-calculate time or frequency domain libraries that are subsequently used during the fitting procedure. This approach allows, as for example in the case of  $Al_{0.5}Ga_{0.5}As$  (discussed below), the calculation of a five-site spectrum, all sites with a quadrupolar parameter distribution, in less than a second on a single 2.3 GHz processor core.

A second requirement for spectrum deconvolution is the choice of a suitable algorithm to the fitting of parameters. If the experimental spectrum contains spectral lines of multiple sites that are not well resolved, the deconvolution model becomes more complex and the number of fit parameters increases. To our knowledge, only one NMR fitting program [7] is able to handle this many parameters, by the use of genetic algorithms. In combination with SIMPSON, this program was used to fit spectra of Vanadium nuclei in haloperoxidases. It, however, is not readily available, and does not separate the fitting from the time consuming simulation step. We introduce the use of evolutionary algorithms [8], of which genetic algorithms are a sub-class, in the *EGdeconv* program. The additional advantage of these algorithms is that they can be efficiently parallelised, leading to an increased computation

\* Corresponding author.

E-mail address: [l.grimminck@science.ru.nl](mailto:l.grimminck@science.ru.nl) (D.L.A.G. Grimminck).

speed. Typical convergence rates involve 100 iterations with a population of 30 trial-solutions. Assuming the case of  $Al_{0.5}Ga_{0.5}As$  (discussed below), this amounts to 35 min for a typical spectrum deconvolution. This calculation time is further reduced by the use of the parallelised code on all available cores (and processors).

Deconvolution of spectra from (partially) disordered materials requires accounting for distributions in interaction parameters, for example by variations in bond angles and/or distances. For the quadrupolar interaction this corresponds to retrieval of the distribution in quadrupolar coupling constant and asymmetry parameter. Czjzek et al. [9] were able to determine the shape of such a distribution for random ordering. By calculating the *efg* tensor components for a large number of random repositionings of surrounding charges, they determined the components become normally distributed. Starting from this multivariate Gaussian distribution of the five independent *efg* tensor components, they could analytically derive a probability density function for the quadrupolar interaction parameters. Le Caër et al. [10,11] extended this model to express the probability density function for a microscopic system with an ordered short-range and disordered long-range surroundings.

Underlying structural variations influence the shape of spectral lines. Deconvolution of the experimental spectrum assisted by these models provides a way to indicate and possibly quantify structural disorder. Currently QuadFit [4] allows the use of the Czjzek model for data-fitting, but no fit-programs are readily available that apply the extended model. Complementary to the microscopic Czjzek model, there is a more macroscopic approach to quantify (dis)order that fixes the relative site abundances given a crystal structure. This fixed set of abundances is related to an order parameter [2,12], of which the value can be varied to range from disordered structure to a targeted crystal structure. Fitting this parameter provides an indication of the material's structure.

The remainder of this paper will describe the Evolutionary Algorithms Serving Your Global Optimisation Improvement Needs Gladly (*EASY-GOING*) deconvolution program, of which the name emphasises its relation to *EASY-GOING DUMBO* homonuclear decoupling [13], that used parts of the same program. We will briefly cover required theory, discuss evolutionary algorithms and the program structure, followed by three case studies and finish with the program's technicalities. In order to learn how to use the program, the reader is referred to the up-to-date manual that is supplied together with the program and example input files at <http://egdeconv.science.ru.nl>.

## 2. Theory

### 2.1. The nuclear quadrupolar interaction

Atomic nuclei that have total angular momentum  $\geq 1$ , possess a none-spherical charge distribution and consequently an electric quadrupolar moment [1]. The Hamiltonian, in units of angular frequency, for the interaction of this moment of the nuclear charge distribution with the *efg* of the environment is given by [6,14]

$$\hat{H}_Q = \sum_i \frac{2\pi C_q^i}{2I_i(2I_i - 1)} \hat{I}_i \cdot \hat{\mathbf{V}}^i \cdot \hat{I}_i. \quad (1)$$

Here  $i$  is the index of the quadrupolar nuclei and  $\hat{I}_i$  and  $I_i$  the nuclear angular momentum vector operator and its quantum number. Furthermore, the reduced *efg* tensor  $\hat{\mathbf{V}}^i$  equals  $\mathbf{V}^i$  divided by its largest principal component, by definition,  $V_{ZZ}$ .  $C_q^i$  is the quadrupolar coupling constant. The general form of  $\mathbf{V}$  and  $C_q$  is

$$C_q = \frac{eQV_{ZZ}}{h} \quad \text{and} \quad V_{\alpha,\beta} = \left. \frac{\partial^2 V}{\partial r_\alpha \partial r_\beta} \right|_{r=0}, \quad (2)$$

with  $V$  the electric potential at the position of the nucleus,  $r_\alpha$  its coordinates for  $\alpha = x, y, z$  and  $eQ$  the nuclear quadrupolar moment. From the form of Eqs. (1) and (2), it follows that  $C_q$  determines the strength of the coupling and additionally provides limits for the position of transition frequencies and, for powdered materials, the width of the spectral lines. For illustrative purposes we rewrite Eq. (1) for the case of a single quadrupolar nucleus in the principal axis frame (PAF) of the *efg* tensor [14]

$$\hat{H}_Q^{PAF} = \frac{2\pi C_q}{4I(2I - 1)} \left\{ 3\hat{I}_z^2 - \hat{I}^2 + \eta (\hat{I}_x^2 - \hat{I}_y^2) \right\}. \quad (3)$$

Here the asymmetry parameter  $\eta$  is defined as

$$\eta = \frac{V_{XX} - V_{YY}}{V_{ZZ}}. \quad (4)$$

It provides symmetry information of the *efg* at the nucleus. The asymmetry parameter can assume values between zero and one, adhering the convention  $|V_{ZZ}| \geq |V_{YY}| \geq |V_{XX}|$ . Combined with the requirement for a traceless *efg* tensor, the quadrupolar interaction is fully described with  $C_q$  and  $\eta$  as parameters.

The dominating Zeeman interaction in NMR makes the form of Eqs. (1) and (3) impractical. Advanced theory for a convenient perturbation treatment of the quadrupolar interaction under magic angle spinning and radio-frequency pulse conditions is extensively treated in literature [5,6,14]. To conclude, it should be noted that the explicit angular dependence of the equations in this section is deliberately not shown. In simulations, the response of the powder is assumed equal to the sum of all crystallite responses. This sum will therefore include all crystallite orientations present, leaving only  $C_q$  and  $\eta$  to parameterise the induction decay/spectrum.

### 2.2. Parameter constraints

Reproducing data by fitting does not always lead to a better understanding of the underlying physics. Essentially the more parameters are involved the more degrees of freedom are available to provide a high quality fit. It is therefore important to search for parameter constraints. These may be obtained from other experimental data, physical limits, etc. *EGdeconv* provides many ways to constrain or couple parameters, which is partly shown below and fully documented in the manual. Another possibility is to impose constraints via (physical) models, as will be introduced in the next two sections.

#### 2.2.1. (Extended) Czjzek distribution

One of the subtleties of the use of quadrupolar NMR to study disordered materials, is the experimental spectrum effectively comprises all local structure variations felt by the type of nucleus under study. The challenge in the analysis is thus to find an inverse mapping of experimental data in terms of structural information, which will in most cases be non-linear. By starting from the notion that structural variations translate into a distribution of the *efg*, a more accessible step is to find the form of this distribution.

Czjzek et al. [9] investigated the *efg* distribution in the case of structural disorder by computer simulations. Tensor components of the *efg* at the centre of randomly repositioned charges, become normally distributed in the limit towards infinite samplings. The resulting probability distribution, as a function of *efg* tensor components, is therefore a multivariate Gaussian. This distribution can be rewritten as a probability density function of quadrupolar interaction parameters  $C_q$  and  $\eta$ , which has the closed-form

$$f(C_q, \eta) = \frac{C_q^{d-1} \eta}{\sqrt{2\pi\sigma^d}} \left( 1 - \frac{\eta^2}{9} \right) \exp \left\{ -\frac{C_q^2}{2\sigma^2} \left( 1 + \frac{\eta^2}{3} \right) \right\}. \quad (5)$$

Here  $\sigma$  is the standard deviation of the multivariate Gaussian probability distribution of the *efg* tensor components and  $d$  is introduced to correspond to the number of independent tensor components [2]. Normalisation of the distribution integral to one, provides the probability of finding a given set of  $C_q$  and  $\eta$ . Later the Czjzek model was revisited and extended by Le Caër and Brand [10,11]. They introduced an ordered (not distributed) contribution to the *efg*, for example originating from a fixed well-defined first coordination sphere. In the *EGdeconv* program we implemented the most general form of the resulting probability density function, given by

$$f(C_q, \eta) \propto \frac{C_q^{d-1}}{\sigma^d} \eta \left(1 - \frac{\eta^2}{9}\right) \int \int \int d\alpha \sin(\beta) d\beta d\gamma. \\ \times \exp \left\{ -\frac{1}{2\sigma^2} \left[ C_{q,0}^2 \left(1 + \frac{\eta_0^2}{3}\right) + C_q^2 \left(1 + \frac{\eta^2}{3}\right) \right. \right. \\ \left. \left. - \frac{2}{\sqrt{3}} C_q C_{q,0} \left( \sqrt{3} a_{11} + \eta_0 a_{15} + \eta \left( a_{51} + \frac{\eta_0 a_{55}}{\sqrt{3}} \right) \right) \right] \right\}. \quad (6)$$

Here  $C_{q,0}$  and  $\eta_0$  are quadrupolar parameters for the local order contribution. The elements  $a_{ij}$  contain the angular dependence of the *efg* tensor components, as defined in [10] and the *EGdeconv* manual. Integration over all angles leaves a distribution as function of quadrupolar interaction parameters that encompasses short-range order and long-range disorder.

### 2.2.2. Order parameter models

In this section we briefly introduce the idea of an order parameter model by treating one specific example. For a more elaborate description we refer to the manual.

In structural investigations of materials, an order parameter is generally introduced to have a measure for the presence of long-range order. This measure is related to the fractional occupation of lattice sites by their preferred atom. While *EGdeconv* has the versatility to couple various parameters and thus allows the implementation of various order parameter models, we have explicitly implemented Copper–Gold and Copper–Platinum ordering in  $A_x B_{1-x} C$  systems as encountered in III–V semiconductors. We immediately introduce the form of the equations for Copper–Gold ordering, meaning a limit to order of alternating (001) planes of A and B atom species

$$p_{C[B_4]} = (1 - r_A)^2 r_B^2 \\ p_{C[AB_3]} = 2r_A(1 - r_A)r_B^2 + 2(1 - r_B)(1 - r_A)^2 r_B \\ p_{C[A_2B_2]} = r_A^2 r_B^2 + (1 - r_B)^2 (1 - r_A)^2 \\ + 4r_A(1 - r_B)r_B(1 - r_A) \quad (7) \\ p_{C[A_3B]} = 2r_A^2(1 - r_B)r_B + 2r_A(1 - r_B)^2(1 - r_A) \\ p_{C[A_4]} = (1 - r_B)^2 r_A^2.$$

The left hand side corresponds to the probability of occurrence for the specific tetrahedral site. The variable  $r_A$  is the fractional occupation number of A atoms present in the correct plane according to the pre-determined order; therefore  $1 - r_A$  is the fraction of not-A atoms present in the same plane. The first equality for  $p_{C[B_4]}$  gives the probability of finding a C atom coordinated by four B atoms. With respect to the Copper–Gold ordered structure this corresponds to two B atoms present in a ‘correct’ and two in an ‘incorrect’ plane. By defining  $r_A = x + 1/2S$  and  $r_B = 1 - x + 1/2S$  we introduce the order parameter  $S$ . Changing  $S$  from zero to a maximum of  $2x$  converts disorder to the maximally achievable Copper–Gold ordering.

## 3. Evolutionary algorithms

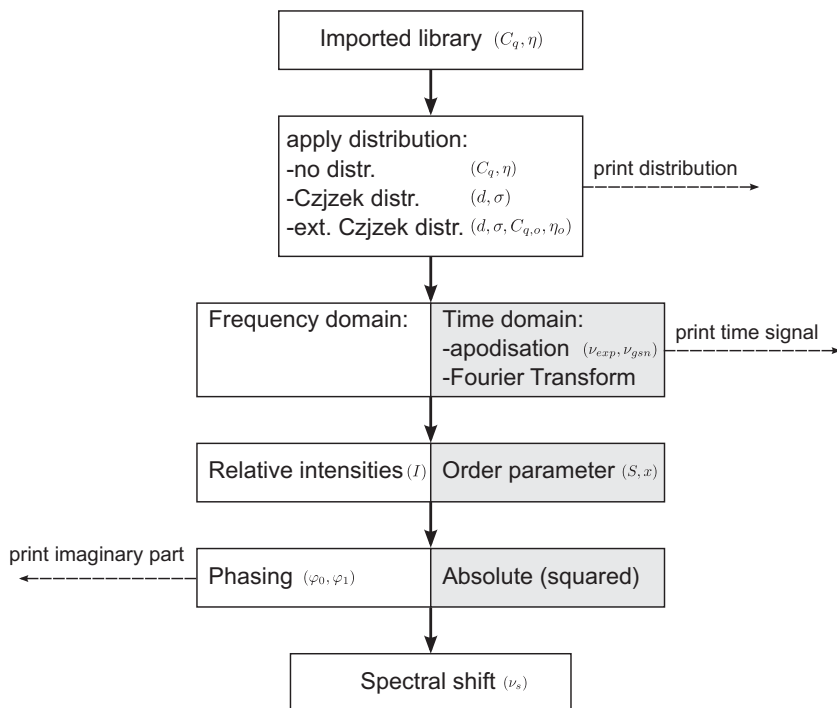
Evolutionary algorithms (EA’s) are based upon the Darwinian theory of a natural selection process occurring by reproduction and mutation of genes in a chromosome leading to only best adapted individuals. In our case, direct translation of these concepts is achieved by accepting vectors with fit parameters as chromosomes and the discrepancy between experimental and theoretical spectrum as selection criterion. EA’s propagate/optimize typically a population of initially, randomly chosen trial solutions in their vector representation. This population-based character, inherently provides more information on search space structure, partly accounts for their global optimisation capabilities. Ultimately it is the algorithm’s mechanism to exploit this search space information that accounts for optimisation performance. *EGdeconv* currently provides the use of three different evolutionary algorithms: a genetic [15], derandomised 2 evolution strategies (DR2-ES) [16] and covariance matrix adaptation evolution strategies (CMA-ES) algorithm [17].

The genetic algorithm includes recombination via parameter vector cross-over’s and applies point mutations to a small part of the population. Evolution strategies algorithms only use overall mutations, the stepsize and direction are determined by strategy parameters. For the DR2-ES algorithm, the mutative steps are determined on inherited strategy parameters from successful parent vectors. The CMA-ES algorithm in addition incorporates parameter covariance information to generate trial solutions. For more in-depth information on the algorithms, we refer to the manual and references given in this section.

## 4. EASY-GOING deconvolution program structure

Here we provide the reader with a brief overview of the program flow, for a more elaborate treatment we refer to the manual at <http://egdeconv.science.ru.nl>. After starting the *EGdeconv* program from the Linux command line, the input file is parsed and the libraries are read into the computer memory. Obviously the amount of computer memory provides a limit to the grid size and resolution of the library, however, *EGdeconv* provides several ways to avoid this limit, for example: by ensuring a library is only present once in the memory, or by allowing to choose the  $[\eta, C_q]$  grid taken from the library or number of data points read per file. Subsequently the selected EA generates parameter vectors that are translated into theoretical spectra. These trial spectra are rated by comparing the experiment to the theory by evaluation of a cost function. There are multiple cost functions defined in *EGdeconv*, for example the use of a weighted innerproduct between the theoretical and experimental spectrum vectors. The EA effectively performs the above in an iterative fashion to reach convergence.

A theoretical spectrum is the sum of subspectra of which each corresponds to a parameter set in the input file and hence a specific chemical site. The program flow to calculate a subspectrum is illustrated in Fig. 1. The figure illustrates in what order the selected library files are processed, and where the various fitting parameters are used. First step is summation of the files according to the distribution model chosen. Then depending on the type of data, time or frequency domain, the result requires additional processing in the form of apodisation and a Fourier transform. Subsequently the intensity of the subspectrum is determined either freely or according to an order parameter model. It is possible to apply a zeroth and first order phase correction to the subspectrum or to use the absolute or absolute squared value of the spectrum. The program flow is ended by positioning the subspectrum. Note that the number of fit parameters per subspectrum can vary between one and effectively ten at maximum. Whether or not *EGdeconv* will



**Fig. 1.** Program flow of *EGdeconv* for the calculation of a subspectrum. The grey-white blocks indicate the choices the user has in its synthesis. For each step, parameters that can be included in the fit are indicated. The definition of part of the parameters ( $d, \sigma, \eta_{(o)}, C_{q(o)}, x$  and  $S$ ) can be found in the Theory section. The exponential/Gaussian apodisation ( $\nu_{exp}/\nu_{gsn}$ ) and zeroth and first order phasing ( $\varphi_0, \varphi_1$ ) are defined in the manual, leaving only the trivial parameters chemical shift ( $\nu_s$ ) and intensity ( $I$ ).

converge depends on a number of factors, for example: the parameter correlations, unique features in the experimental spectrum and of course the accuracy of the simulated library files.

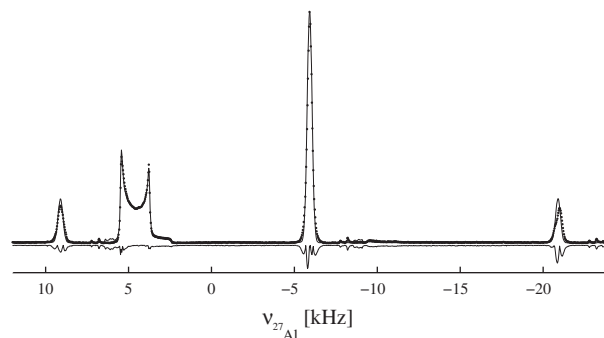
Fig. 1 also illustrates that *EGdeconv* can print extra output. This feature provides the user with more insight in what happens during the different stages in the calculation. The amount of output is governed by appropriately setting keywords in the input file.

## 5. Examples

In this section three applications of the *EGdeconv* program are presented. They represent the program's versatility, by the use of second-party software, in this work SIMPSON [5], to handle different experimental conditions, for example: static or magic angle spinning (MAS), free induction decay or Hahn echo acquisition, etc. Additionally, we show lineshape analysis with or without (extended) Czjzek distributions. Input files for these examples are provided with the installation package.

### 5.1. $^{27}\text{Al}$ MAS NMR spectrum of yttrium aluminium garnet

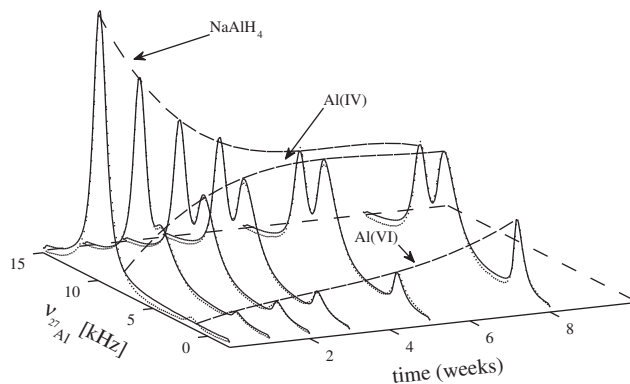
Fig. 2 shows a hard pulse  $^{27}\text{Al}$  MAS NMR spectrum of YAG ( $\text{Y}_3\text{Al}_5\text{O}_{12}$ ) at 15 kHz MAS. YAG has a garnet type structure with Al occurring in a tetrahedral and an octahedral coordination in 3:2 ratio. The spectrum of powdered YAG was measured and fully characterised by Massiot et al. in 1990 [18]. Combination of MAS conditions and the well-defined quadrupolar interactions, indicated by the sharp lineshapes, make this a suitable benchmark study for *EGdeconv*. Massiot et al. established interaction constants for the tetrahedrally ( $\sim 6$  MHz,  $\eta = 0$ ) and octahedrally ( $\sim 0.6$  MHz,  $\eta = 0$ ) coordinated sites with a relative abundance ( $\sim 3:2$ ). Given the results mentioned in the caption of Fig. 2, *EGdeconv* efficiently reaffirmed these findings.



**Fig. 2.** Experimental (dotted line) and theoretical (continuous line)  $^{27}\text{Al}$  spectrum of YAG together with the negative absolute value of the difference. The spectrum consists of a broad resonance of the tetrahedral site in addition to two spinning side bands and the central transition of the octahedral site. Five independent fits with  $[\eta-C_q]$  grids [0.00:0.01:0.10-5.0:0.1:7.0] and [0.00:0.01:0.10-0.3:0.1:1.0] obtained  $C_q$  values of 6.1 and 0.6 MHz and  $\eta$  values of 0.04 and  $0.07 \pm 0.02$ , respectively, only standard deviations larger than the library stepsize are noted. The relative intensity of the respective sites, was determined 1 : ( $0.69 \pm 0.01$ ).

### 5.2. Oxidation of sodium aluminium hydride ( $\text{NaAlH}_4$ )

Fig. 3 shows the results of fitting  $^{27}\text{Al}$  MAS spectra of  $\text{NaAlH}_4$ 's oxidation process as a function of time. All information provided here is based on [19] that contains references and more details about the data analysis. The  $\text{NaAlH}_4$  sites in the material have a well defined lattice without significant variations. When exposed to air, structurally disordered tetrahedrally and octahedrally coordinated oxide sites will form. From Multiple Quantum MAS experiments it was determined that the quadrupolar interaction of the tetrahedral (Al(IV)) site is distributed. Therefore the lineshape of this site was fitted with a Czjzek distribution that reproduces the asymmetric lineshape, as shown in Fig. 3. The octahedral site (Al(VI)) contained in addition to a distribution in quadrupolar



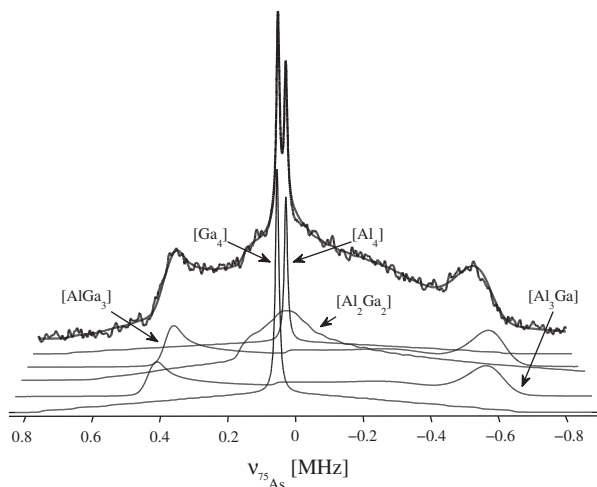
**Fig. 3.** Oxidation of  $\text{NaAlH}_4$  as a function of time. Experimental (dotted line) and theoretical spectra (contuous line) are overlaid and shown as a function of time. The dashed lines roughly connecting the peaks are qualitative guides to the eye. They emphasize the conversion of  $\text{NaAlH}_4$  in tetrahedral  $\text{Al(IV)}$  and octahedral  $\text{Al(VI)}$  oxide sites. For the fits a  $[\eta-C_q]$  grid of [0.00:0.05:1.00–0.5:0.5:15.0] was used.

parameters also line broadening from dipole–dipole couplings, this influence was mimicked by additional line broadening. As the  $\text{NaAlH}_4$  has a well-defined crystal structure this line was described without a distribution in quadrupole parameters. Additional Gaussian line broadening accounted for the dipolar broadening as the spectra were obtained without proton decoupling.

The aim of the spectral deconvolution here was to quantify the relative intensities of the various lines as a function of time. This allows to monitor the interconversion of the different components during oxidation as shown in Fig. 3, and discussed in detail in [19].

### 5.3. Static hahn echo spectrum of aluminium gallium arsenide ( $\text{Al}_x\text{Ga}_{1-x}\text{As}$ )

Fig. 4 shows a  $^{75}\text{As}$  Hahn echo spectrum of powdered  $\text{Al}_{0.489}\text{Ga}_{0.511}\text{As}$ , one form of the widely used and well-characterised  $\text{Al-GaAs}$  semi-conductor. Solid-state NMR was used to investigate the possible order or disorder in the occupation of the  $\text{Al/Ga}$  sites.



**Fig. 4.** Experimental (thick dotted line) and theoretical (thick continuous line) Hahn echo spectrum of  $\text{Al}_{0.489}\text{Ga}_{0.511}\text{As}$ , deconvoluted in terms of subspectra corresponding to tetrahedral aluminium–gallium sites. The  $[\text{Al}_4]$  and  $[\text{Ga}_4]$  sites were fitted with a Czjzek distribution, the remaining sites required an extended Czjzek distribution to account for the ordered  $efg$  contribution in the first coordination sphere. Subspectrum intensities were fitted with respect to one subspectrum, but independently. Three libraries with  $[\eta-C_q]$  grids of [0.00:0.05:1.00–0.0:0.2:3.2] ( $[\text{Al}_4]$  and  $[\text{Ga}_4]$ ), [0.00:0.01:0.10–31.0:0.1:35.0] ( $[\text{Al}_3\text{Ga}]$  and  $[\text{AlGa}_3]$ ) and [0.85:0.01:1.00–28.0:0.1:35.0] ( $[\text{Al}_2\text{Ga}_2]$ ) were used.

The analysis of this spectrum is published in [2] and was the main driving force behind the development of *EGdeconv*.

With a spectral width of 2.5 MHz and sites with small and large quadrupolar coupling constants present, this example is particularly dependent on accurately simulated excitation efficiencies. As the figure shows, the spectrum can be decomposed in terms of tetrahedral aluminium–gallium coordinations of the arsenic atoms. The centred narrow lines correspond to the  $[\text{Al}_4]$  and  $[\text{Ga}_4]$  coordinations of the arsenic atoms, these sites have such a high symmetry that the  $efg$  due to the first atomic coordination shell is absent, leading to a small quadrupolar interaction strength caused by disorder in the higher coordination spheres. With respect to excitation efficiencies it should be noted that for these sites both central and satellite transitions are of importance. Due to a negligible  $efg$  contribution of the first atom coordination shell, these sites are specific probes of disorder in higher order coordination spheres. This justifies the use of a Czjzek distribution to calculate the lineshapes. The symmetry of the remaining three sites causes a large  $efg$  contribution from the first coordination shell. The excitation efficiency of these sites is mainly selective to the central  $\langle 1/2, -1/2 \rangle$  transition. To fit these sites an extended Czjzek distribution was applied, for which the  $\sigma$  parameter was set to the average of the  $[\text{Al}_4]$  and  $[\text{Ga}_4]$  Czjzek distribution, since the latter are probes for long range disorder. The last step in the analysis is the use of an order parameter to relate the site intensities. The resulting fit, shown in [2] and available in the installation package, matches less than that of Fig. 4, but provides a value of  $S$  close to zero, indicating the expected structural disorder. An in-depth study of this disorder supported by DFT calculations reproducing the fitted distributions is described in [2].

## 6. Conclusions and outlook

We presented *EASY-GOING deconvolution*, a program for fast and accurate fitting of one-dimensional solid-state quadrupolar NMR spectra of powdered materials. The use of second-party software to calculate libraries as a function of quadrupolar parameters allows accurate simulation of the experiment and low computational costs for theoretical spectrum synthesis during the fitting. This furthermore allows a straightforward extension of the program to fit other NMR interactions, such as chemical shift anisotropy and combinations thereof. The program is provided for 64-bit Linux operating systems and supports efficient parallel computing due to the nature of the incorporated evolutionary algorithms that can handle many-parameter fits. The program includes the (extended) Czjzek distribution and order parameter model to indicate and quantify structural (dis)order in materials. With three examples we showed *EGdeconv*'s versatility to manage different experiment types and quadrupolar parameter distributions. In the future the program will be extended to include the fitting of single crystal data and 2D spectra.

## 7. Experimental

### 7.1. Experimental data

The  $^{27}\text{Al}$  spectrum of Fig. 2 was measured on a 14.1 T Chemagnetics Infinity NMR spectrometer, using a 2.5 mm double resonance probe. The experiment consisted of a single 0.2  $\mu\text{s}$  hard pulse at 214 kHz rf field strength, while magic angle spinning the sample at a frequency of 15 kHz.

Experimental parameters for the spectra in Fig. 3 were a hard pulse excitation of 0.2  $\mu\text{s}$  and 160 kHz effective rf field strength at 15 kHz MAS at 9.1 T external field. The Hahn echo experiments on  $\text{AlGaAs}$  in Fig. 4 were performed at 18.8 T external field with an

effective rf field of 625 kHz. The first delay was set to 175  $\mu$ s and acquisition started 5  $\mu$ s after the  $\pi$  pulse. For more information on the latter two experiments we refer to [19,2], respectively.

## 7.2. How to use EASY-GOING deconvolution

This section covers the practical aspects of *EGdeconv*. Introduction of input for *EGdeconv* adheres the convention that  $\langle \dots \rangle$  input is required and [...] is optional. The words, units and abbreviations replacing the ... help to interpret the meaning of the variable as to what should be filled in. Please note that for the full description of all topics in this section we refer to the manual at <http://egdeconv.science.ru.nl>.

### 7.2.1. Installation requirements

The *EGdeconv* program requires a 64-bit Linux operating system. The installation package includes the *EGdeconv* and OpenMPI binaries. OpenMPI is responsible for the parallel computing. The program has been tested on various Linux distributions: Fedora 9, 10 and 11, Ubuntu 10.04 (LTS) and openSUSE 11.3. Installation will proceed via a shell script as depicted on our website.

### 7.2.2. Experimental data

*EGdeconv* requires the one-dimensional experimental spectrum in a two-column ASCII file with the number of complex spectrum points a power of two. The first column should be the spectral frequency and the second, the real or absolute value (squared) of the spectrum.

### 7.2.3. Libraries

Before attempting to fit the experimental data, the user should calculate a library of time or frequency domain data as a function of  $\eta$  and  $C_q$  for the appropriate experimental conditions. This can be done with any preferred simulation program, for example SIMPSON [5] and SPINEVOLUTION [6], as long as the libraries adhere to the format given in this section. The filename format should be  $\langle \text{sitename} \rangle - \langle C_q(\text{MHz}) \rangle - \langle \eta \rangle . \langle \text{arb. file extension} \rangle$

for example A1O4–30.25–0.01.fid. The currently supported internal file format is that of SIMPSON. The first five lines contain, in plain text, the keywords: SIMP, NP =  $\langle \# \text{points} \rangle$ , SW =  $\langle \text{MHz} \rangle$ , TYPE = -FID/SPE and DATA. This information is used by *EGdeconv*. The rest of the file consists of the real and imaginary parts of the calculated data in the first and second column, respectively, finalised by the keyword END. *EGdeconv* allows to regulate the size of the actual  $(C_q, \eta)$ -grid used for the fitting procedure, via the keyword 'speciesX' as will be discussed below. Note that spectrum files should have the same number of points as the experiment, for the time domain data this is irrelevant. Examples of libraries are included in the installation package.

### 7.2.4. Input file

The *EGdeconv* installation package contains a number of example input files. Input files for the *EGdeconv* binary consist basically of plain text containing lines with either: comments (starting with #), keywords

$\langle \text{keyword string} \rangle = \langle \text{value(s)} \rangle / \langle \text{string} \rangle$

or parameter settings.

parameter =  $\langle n \rangle \langle \text{name} \rangle \langle \text{start} \rangle [ \langle \text{lower} \rangle [ \langle \text{upper} \rangle ] ] [ \langle \text{coupling} \rangle ]$   
 $\times [ \langle \text{nset/num} \rangle ]$

There are numerous keywords, all tabulated in the manual, enabling the choice of: EA, extra output, special weighting or removal of parts of the spectrum, etc. a few examples are

```
#relative path
experimental spectrum = expdata/spec.dat
#absolute path to library
species1 = /home/user/depart_lib/AI27
#single calc with start parameters
testcalculation = 1
```

Parameter lines contain, according to the format given above, parameter: reference number, (arbitrary) name, fixed value for single spectrum calculation or during fitting in absence of boundaries, upper and lower boundaries for fitting and type of inter/intra parameter set couplings. A few examples

```
#fix value during fit
parameter = 1 d 1
#fit between 3 and 5 MHz
parameter = 2 sigma 4 3 5
#value equal to par 3 set 1
parameter = 3 qcc0 30 E 1
```

It is important to place parameter lines in order of increasing reference number; the number indicates the parameter's identity for *EGdeconv*. The *EGdeconv* parser will consider the subsequent encounter of a lower reference number as the start of a new parameter set. Parameter sets correspond to individual sites and their (sub)spectrum.

One keyword named 'species X' requires special attention. This is the link between the input file and the calculated libraries. Creative usage of this keyword can save both computation time and memory. The keyword with all options present is given by

species(X) =  $\langle \text{path} \rangle [ \langle \eta_b \rangle [ \langle \eta_e \rangle [ \langle \eta_s \rangle [ \langle C_{q,b} \rangle [ \langle C_{q,e} \rangle [ \langle C_{q,s} \rangle ] ] ] ] ] ] ] ]$

For a different library the value of X, starting from one, should increase by one. The subscripts: *b*(egin), *e*(nd), and *s*(tep), indicate the precise grid in  $\eta$  and  $C_q(\text{MHz})$  to use from the library located at  $\langle \text{path} \rangle$ . If no grid is specified, *EGdeconv* will load the entire library, unless the keyword is not linked to a parameter set.

In addition to the keyword 'speciesX' there is also an equally named parameter, currently reference number 199. The start value of this parameter relates directly to X and the corresponding library. This system allows multiple parameter sets to use the same library while it is present only once in the computer memory.

A second special parameter, currently number 198, regulates the quadrupolar parameter distribution model used to calculate the (sub)spectrum for the parameter set. The name of the parameter should be either: *no\_distr*, *czj\_distr* or *ext\_czj\_distr* which are self explanatory. The choice of distribution model assigns the meaning of parameter reference numbers one to four to respectively  $(\eta, C_q)$ ,  $(d, \sigma)$  and  $(d, \sigma, C_{q,o}, \eta_o)$ . More information on the available parameters and their couplings is available in the *EGdeconv* manual. Working examples are also supplied with the installation package.

### 7.2.5. Program output

*EGdeconv* facilitates various outputs during, and at the end of the fitting procedure or calculation using only parameter start values for a test spectrum. During the fitting, a file is written and updated to summarise the current best fit parameters and the cost function value. At the end of the procedure a new file is created that contains the columns: frequency, best fit spectrum, experimental data and subspectra. Additionally a final file is made that summarises the results of the fit. We refer to the manual for a complete description of the (optional) output.

## Acknowledgements

The authors thank Paulus J. Knijn, Ernst R.H. van Eck and Margriet W.H. Verkuijlen for providing experimental spectra for analysis using *EGdeconv*. We thank Gerrit C.J. Groenenboom for useful discussions on theoretical considerations and testing of the program. The Netherlands Organisation for Scientific Research (NWO) are acknowledged for their financial support of the solid-state NMR facility for advanced material science.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmr.2011.04.009](https://doi.org/10.1016/j.jmr.2011.04.009).

## References

- [1] A. Abragam, *The Principles of Nuclear Magnetism*, The International Series of Monographs on Physics, Oxford University Press, 1967.
- [2] P.J. Knijn, P.J.M. van Bentum, E.R.H. van Eck, C. Fang, D.L.A.G. Grimminck, R.A. de Groot, R.W.A. Havenith, M. Marsman, W.L. Meerts, G.A. de Wijs, A.P.M. Kentgens, A solid-state NMR and DFT study of compositional modulations in  $\text{Al}_x\text{Ga}_{1-x}\text{As}$ , *Phys. Chem. Chem. Phys.* 12 (2010) 11517–11535.
- [3] D. Massiot, F. Fayon, M. Capron, I. King, S. Le Calvé, B. Alonso, J. Durand, B. Bujoli, Z. Gan, G. Hoatson, Modelling one- and two-dimensional solid-state NMR spectra, *Magn. Reson. Chem.* 40 (2002) 70–76.
- [4] T.F. Kemp, M.E. Smith, QuadFit – a new cross-platform computer program for simulation of NMR line shapes from solids with distributions of interaction parameters, *Solid State Nucl. Magn. Reson.* 35 (2009) 243–252.
- [5] M. Bak, J.T. Rasmussen, N.C. Nielsen, SIMPSON: A general simulation program for solid-state NMR spectroscopy, *J. Magn. Reson.* 147 (2000) 296–330.
- [6] M. Veshkort, R.G. Griffin, SPINEVOLUTION: a powerful tool for the simulation of solid and liquid state NMR experiments, *J. Magn. Reson.* 178 (2006) 248–282.
- [7] M. Wächtler, A. Schweitzer, T. Gutmann, H. Breitzke, G. Buntkowsky, Efficient analysis of  $^{51}\text{V}$  solid-state MAS NMR spectra using genetic algorithms, *Solid State Nucl. Magn. Reson.* 35 (2009) 37–48.
- [8] W.L. Meerts, M. Schmitt, Application of genetic algorithms in automated assignments of high-resolution spectra, *Int. Rev. Phys. Chem.* 25 (2006) 353–406.
- [9] G. Czjzek, J. Fink, F. Götz, H. Schmidt, J.M.D. Coey, J.P. Rebouillat, A. Liénard, Atomic coordination and the distribution of electric field gradients in amorphous solids, *Phys. Rev. B* 23 (1981) 2513–2530.
- [10] G. Le Caër, R.A. Brand, General models for the distributions of electric field gradients in disordered solids, *J. Phys.: Condens. Matter* 10 (1998) 10715–10774.
- [11] G. Le Caër, B. Bureau, D. Massiot, An extension of the czjzek model for the distributions of electric field gradients in disordered solids and an application to NMR spectra of  $^{71}\text{Ga}$  in chalcogenide glasses, *J. Phys.: Condens. Matter* 22 (2010) 065402.
- [12] R. Tycko, G. Dabbagh, S.R. Kurtz, J.P. Goral, Quantitative study of atomic ordering in  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$  thin films by  $^{31}\text{P}$  nuclear magnetic resonance, *Phys. Rev. B* 45 (1992) 13452–13457.
- [13] D.L.A.G. Grimminck, S.K. Vasa, A. Brinkmann, W.L. Meerts, A.P.M. Kentgens, EASY-GOING DUMBO On-Spectrometer Optimisation of Phase Modulated Homonuclear Decoupling Sequences in Solid-State NMR, *Chem. Phys. Lett.*, in press. [doi.org/10.1016/j.cplett.2011.04.079](https://doi.org/10.1016/j.cplett.2011.04.079).
- [14] M.J. Duer, *Introduction to Solid-State NMR Spectroscopy*, Blackwell, 2004.
- [15] J.A. Hageman, R. Wehrens, R. de Gelder, W.L. Meerts, L.M.C. Buydens, Direct determination of molecular constants from rovibronic spectra with genetic algorithms, *J. Chem. Phys.* 113 (2000) 7955–7962.
- [16] A. Ostermeier, A. Gawelcyk, N. Hansen, Step-size adaptation based on non-local use of selection information, in: Y. Davidor, H.P. Schwefel, R. Männer (Eds.), *Parallel Problem Solving from Nature, PPSN III*, Springer, Berlin/Heidelberg, 1994.
- [17] N. Hansen, A. Ostermeier, Completely derandomized self-adaptation in evolution strategies, *Evol. Comput.* 9 (2001) 159–195.
- [18] D. Massiot, C. Bessada, J.P. Coutures, F. Taulelle, A quantitative study of  $^{27}\text{Al}$  MAS NMR in crystalline YAG, *J. Magn. Reson.* 90 (1990) 231–242.
- [19] M.H.W. Verkuijlen, R. de Gelder, P.J.M. van Bentum, A.P.M. Kentgens, The oxidation products of  $\text{NaAlH}_4$  studied by solid-state NMR and X-ray diffraction, *J. Phys. Chem. C* (2011). [Dx.doi.org/10.1021/jp109927a](https://doi.org/10.1021/jp109927a).