



Cite this: *Phys. Chem. Chem. Phys.*,  
2016, **18**, 26412

Received 5th August 2016,  
Accepted 23rd August 2016

DOI: 10.1039/c6cp05437d

www.rsc.org/pccp

## Pseudocontact shifts from mobile spin labels†

Elizaveta A. Suturina\* and Ilya Kuprov

This paper presents a detailed analysis of the pseudocontact shift (PCS) field induced by a mobile spin label that is viewed as a probability density distribution with an associated effective magnetic susceptibility anisotropy. It is demonstrated that non-spherically symmetric density can lead to significant deviations from the commonly used point dipole approximation for the PCS. Analytical and numerical solutions are presented for the general partial differential equation that describes the non-point case. It is also demonstrated that it is possible, with some reasonable approximations, to reconstruct paramagnetic centre probability distributions from the experimental PCS data.

### 1. Introduction

Pseudocontact shift (PCS) is an additional contribution to the nuclear chemical shift caused by the presence of a paramagnetic centre in close proximity to the nucleus in question.<sup>1,2</sup> PCS is very well researched and is widely used as a source of structural restraints for paramagnetic metalloproteins.<sup>3–7</sup> Even when a protein is not naturally paramagnetic, the commonly occurring calcium, magnesium and zinc binding sites would usually coordinate a lanthanide well.<sup>8</sup> In combination with artificially introduced lanthanide-containing tags, PCS is also used to determine relative orientations of protein domains.<sup>9</sup> In magnetic resonance imaging it is useful as a reporter for local pH and oxidation potential.<sup>10,11</sup>

General equations describing chemical shielding, obtained by an assiduously systematic application of perturbation theory, are due to Ramsey.<sup>12,13</sup> The first paper dealing with a point dipole approximation for chemical shift was published by McConnell<sup>14</sup> in 1957 – he noted that shielding by sufficiently distant electrons could be expressed *via* an effective magnetic susceptibility tensor; for paramagnetic molecules this tensor is a function of the spin Hamiltonian parameters of the paramagnetic centre.<sup>15</sup> Analytical treatments for specific classes of *d*- and *f*-transition metal complexes using ligand field theory have been reported by Bleaney,<sup>16</sup> Golding,<sup>17</sup> and Stiles.<sup>18–21</sup>

Modern quantum chemistry defines the paramagnetic shift as the Frobenius inner product between the nuclear hyperfine coupling tensor and the magnetic susceptibility tensor.<sup>22,23</sup> Both parameters are difficult to compute because they have contributions from spin–orbit coupling and often require

non-perturbative treatment of relativistic effects within multi-configurational *ab initio* methods,<sup>24,25</sup> as well as conformational averaging. Accurate quantum chemical calculations are therefore limited to a few dozen atoms.

Structural biologists couldn't care less – most nuclei in macromolecules are far enough away from the paramagnetic centre for McConnell's approximation to be accurate<sup>2,26</sup> and the resulting formula for the pseudocontact shift produced by a point source with a given magnetic susceptibility anisotropy has been of great service to protein structure and dynamics research over the last 30 years.<sup>8,27</sup> Excellent software packages exist that make PCS analysis in proteins and nucleic acids straightforward and informative.<sup>5,28</sup>

There remains one important unsolved problem: pseudocontact shift prediction and analysis in large systems that feature fast conformational mobility of the paramagnetic centre. For such systems the point dipole approximation is no longer valid at short distances, and quantum chemical calculations are prohibitively expensive. For this reason, PCS measurements in close proximity to the tag are often excluded from the analysis because they are not expected to fit the point dipole formula.<sup>29</sup>

In this paper, we introduce an analytical approach based on the recently published partial differential equation for PCS<sup>30</sup> that views the paramagnetic centre as a probability density distribution in three dimensions. This approach clarifies the key features of that density that affect PCS. It may also be used to recover the distribution itself from the experimental PCS data.

### 2. Pseudocontact shift from a point paramagnetic centre

The point source formula for PCS is best derived using a classical physics argument. Placed in an external magnetic field  $\mathbf{B}_0$ , a

School of Chemistry, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, UK. E-mail: e.suturina@soton.ac.uk

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6cp05437d



paramagnetic centre would acquire the following magnetic dipole moment:

$$\boldsymbol{\mu}_e = \chi \cdot \mathbf{B}_0 / \mu_0 \quad (1)$$

where  $\mu_0$  is the vacuum permeability and  $\chi$  is the magnetic susceptibility tensor.<sup>26</sup> This linear response assumption is valid for an ensemble of non-interacting paramagnetic centres when  $|\boldsymbol{\mu}_e^T \cdot \mathbf{B}_0| \ll kT$  – true for most metal–organic systems at room temperature. For a point centre, the induced dipole creates the following magnetic field at the relative position  $\mathbf{r}$ :

$$\mathbf{B}_1 = \frac{3\mu_0}{4\pi r^3} \mathbf{D}(\hat{\mathbf{r}}) \cdot \boldsymbol{\mu}_e, \quad (2)$$

where the dipolar matrix is:

$$\mathbf{D}(\hat{\mathbf{r}}) = \frac{1}{r^2} \begin{bmatrix} x^2 - \frac{r^2}{3} & xy & xz \\ xy & y^2 - \frac{r^2}{3} & yz \\ xz & yz & z^2 - \frac{r^2}{3} \end{bmatrix} \quad (3)$$

It is easy to demonstrate that the dipolar matrix only depends on the direction of the position vector  $\hat{\mathbf{r}} = \{\theta, \phi\}$  and does not depend on its length  $r = |\mathbf{r}|$ . This provides a clean separation of coordinates that will be useful below.

The change in the energy of a nuclear magnetic moment  $\boldsymbol{\mu}_n$  produced by placing it at the position  $\mathbf{r}$  relative to the paramagnetic centre would be

$$\Delta E = -\boldsymbol{\mu}_n^T \cdot \mathbf{B}_1 \quad (4)$$

The associated chemical shift tensor is the second derivative of  $\Delta E$  with respect to  $\boldsymbol{\mu}_n$  and  $\mathbf{B}_0$ :<sup>12,31</sup>

$$\boldsymbol{\sigma}^{\text{point}} = -\frac{\partial^2 E}{\partial \boldsymbol{\mu}_n^T \partial \mathbf{B}_0} = \frac{3}{4\pi r^3} \mathbf{D}(\hat{\mathbf{r}}) \cdot \chi \quad (5)$$

The isotropic average of this tensor is the familiar point dipole expression for the pseudocontact shift:<sup>14,26</sup>

$$\begin{aligned} \sigma^{\text{point}} &= \frac{1}{3} \text{Tr}[\boldsymbol{\sigma}^{\text{point}}] = \frac{1}{4\pi r^3} \text{Tr}[\mathbf{D}(\hat{\mathbf{r}}) \cdot \chi] \\ &= \frac{1}{4\pi r^3} \left( \frac{\mathbf{r}^T \cdot \chi \cdot \mathbf{r}}{\mathbf{r}^T \cdot \mathbf{r}} - \frac{1}{3} \text{Tr}(\chi) \right) \end{aligned} \quad (6)$$

that may also be written *via* second-rank spherical harmonics:

$$\sigma^{\text{point}} = \frac{1}{4\pi r^3} \sum_{m=-2}^2 \chi_m Y_2^m(\hat{\mathbf{r}}) \quad (7)$$

where  $\chi_m$  are the irreducible spherical components of  $\chi$ :

$$\begin{aligned} \chi_{\pm 2} &= +\sqrt{\frac{2\pi}{15}} (\chi_{XX} - \chi_{YY} \mp i(\chi_{XY} + \chi_{YX})) \\ \chi_{\pm 1} &= -\sqrt{\frac{2\pi}{15}} (\chi_{XZ} + \chi_{ZX} \mp i(\chi_{YZ} + \chi_{ZY})) \\ \chi_0 &= +\sqrt{\frac{4\pi}{45}} (2\chi_{ZZ} - (\chi_{XX} + \chi_{YY})) \end{aligned} \quad (8)$$

It is important to note that the isotropic part and the first spherical rank component of the magnetic susceptibility tensor do not enter the equation for PCS. The five irreducible spherical tensor parameters in eqn (8) may also be expressed as axially and rhombicity, along with the three parameters (*e.g.* Euler angles) specifying the orientation of the principal axis frame.

### 3. Pseudocontact shift from a distributed paramagnetic centre

A less well explored situation is when the paramagnetic centre is distributed with some probability density  $\rho(\mathbf{r}_e)$  within the molecular structure. In such a situation, the magnetic susceptibility tensor would also in general be position-dependent. The integration of the point PCS expression in eqn (5) over the probability density produces the following expression for the dipolar shift tensor at position  $\mathbf{r}$ :

$$\boldsymbol{\sigma}(\mathbf{r}) = \frac{3}{4\pi} \int \frac{\mathbf{D}(\mathbf{r}_e - \mathbf{r})}{|\mathbf{r}_e - \mathbf{r}|^3} \chi(\mathbf{r}_e) \rho(\mathbf{r}_e) d^3 \mathbf{r}_e \quad (9)$$

This integral is a convolution of the dipolar matrix  $\mathbf{D}$  divided by the cube of the distance with the product of susceptibility tensor and probability density:

$$\boldsymbol{\sigma}(\mathbf{r}) = \frac{3}{4\pi} \frac{\mathbf{D}(\mathbf{r})}{r^3} * [\chi(\mathbf{r}) \rho(\mathbf{r})] \quad (10)$$

The simplest way to proceed is to use the Fourier transform because convolution is equivalent to multiplication in the  $k$ -space, and the Fourier transform of the dipolar matrix is very simple:

$$\text{FT}_+ \left[ \frac{3}{4\pi r^3} \mathbf{D}(\hat{\mathbf{r}}) \right] = -\mathbf{D}(\hat{\mathbf{k}}) \quad (11)$$

where  $\hat{\mathbf{k}}$  is the angular part of the  $k$ -space vector  $\mathbf{k}$ . Another useful property of the dipolar matrix is that the inverse Fourier transform of its product with a function in  $k$ -space can be expressed as an action of the differential operator in real space:

$$\boldsymbol{\sigma}(\mathbf{r}) = -\text{FT}_- [\mathbf{D}(\hat{\mathbf{k}}) \hat{\mathbf{g}}(\mathbf{k})] = -\mathbf{S} \mathbf{g}(\mathbf{r}) \quad (12)$$

where  $\hat{\mathbf{g}}(\mathbf{k})$  is the Fourier transform of  $\mathbf{g}(\mathbf{r}) = \chi(\mathbf{r}) \rho(\mathbf{r})$  and the differential operator  $\mathbf{S}$  has the following form:

$$\mathbf{S} = \frac{\vec{\nabla} \otimes \vec{\nabla}^T}{\vec{\nabla}^T \cdot \vec{\nabla}} - \frac{1}{3} = \frac{1}{\nabla^2} \begin{bmatrix} \partial_{xx} - \frac{\nabla^2}{3} & \partial_{xy} & \partial_{xz} \\ \partial_{yx} & \partial_{yy} - \frac{\nabla^2}{3} & \partial_{yz} \\ \partial_{zx} & \partial_{zy} & \partial_{zz} - \frac{\nabla^2}{3} \end{bmatrix} \quad (13)$$

This is a reciprocal equation to eqn (3) – multiplication by a vector in Fourier space is equivalent to taking a gradient in real space and division by  $r^2$  in real space is equivalent to the inverse Laplacian in  $k$ -space.

In the case where both the probability density and the magnetic susceptibility tensor are position-dependent, this operator acts



on their product and the following general expression is obtained for the matrix elements of the paramagnetic shift tensor:

$$\sigma_{ij}(\mathbf{r}) = -\frac{1}{\sqrt{2}} \sum_k \partial_{ik} [\chi_{kj}(\mathbf{r})\rho(\mathbf{r})] + \frac{\delta_{ij}}{3} \chi_{ij}(\mathbf{r})\rho(\mathbf{r}) \quad (14)$$

The second term on the right hand side effectively subtracts the Fermi contact part of the full paramagnetic shift, ensuring that PCS does not depend on  $\text{Tr}(\chi)$  in the same way as it happens in the point model.

Eqn (14) simplifies significantly under the assumption that the susceptibility tensor is the same at all locations. After the isotropic average is taken, the result is:

$$\sigma(\mathbf{r}) = -\frac{1}{3} \left[ \frac{\vec{\nabla}^T \cdot \chi \cdot \vec{\nabla}}{\vec{\nabla}^T \cdot \vec{\nabla}} - \frac{1}{3} \text{Tr}(\chi) \right] \rho(\mathbf{r}) \quad (15)$$

In the Fourier space this equation can be written as

$$\hat{\sigma}(\mathbf{k}) = -\frac{1}{3} \left[ \frac{\mathbf{k}^T \cdot \chi \cdot \mathbf{k}}{\mathbf{k} \cdot \mathbf{k}^T} - \frac{1}{3} \text{Tr}(\chi) \right] \hat{\rho}(\mathbf{k}) \quad (16)$$

where  $\hat{\rho}(\mathbf{k})$  is the Fourier transform of  $\rho(\mathbf{r})$  and  $\hat{\sigma}(\mathbf{k})$  is the Fourier transform of  $\sigma(\mathbf{r})$ .

In the derivation presented above,  $\rho(\mathbf{r})$  is the statistical probability density of the paramagnetic centre, but one can make an approximate parallel here with the quantum mechanical spin density. The dipolar part of the hyperfine coupling at the nuclear position  $\mathbf{r}$  for a given spin density  $\rho^{\text{spin}}(\mathbf{r}_e)$  is

$$\mathbf{A}^{\text{dip}}(\mathbf{r}) = \frac{3\mu_0\gamma_e\gamma_n\hbar}{4\pi} \int \frac{\mathbf{D}(\mathbf{r}_e - \mathbf{r})}{|\mathbf{r}_e - \mathbf{r}|^3} \rho^{\text{spin}}(\mathbf{r}_e) d^3\mathbf{r}_e \quad (17)$$

and the PCS computed *ab initio* leads to the same equation but with the spin density

$$\sigma(\mathbf{r}) = -\frac{1}{3} \frac{\text{Tr}[\mathbf{A}^{\text{dip}}(\mathbf{r}) \cdot \chi]}{\hbar\mu_0\gamma_e\gamma_n}, \quad (18)$$

meaning that the dipolar hyperfine coupling tensor field, viewed as an integral over the spin density, can also be expressed using the differential operator from eqn (13):

$$\mathbf{A}^{\text{dip}}(\mathbf{r}) = -\hbar\mu_0\gamma_e\gamma_n \mathbf{S} \rho^{\text{spin}}(\mathbf{r}). \quad (19)$$

It must be noted that eqn (18) only accounts for the dipolar part of the hyperfine coupling and does not include the orbital contribution, which is important in the immediate vicinity of heavy ions.<sup>32,33</sup>

## 4. Analytical solution to the direct problem

The “direct” problem will be defined here as the task of calculating the PCS from a given magnetic susceptibility tensor and a given probability distribution of the paramagnetic centre at the specified nuclear coordinates. The general case with a position-dependent magnetic susceptibility tensor is described by eqn (14). It is clear from the form of eqn (14) that, in order to disentangle the effects of the two spatial functions (density and

susceptibility), we need to know *a priori* at least one of them. Below we analyse the special case where the magnetic susceptibility tensor is the same at every point of the probability density in eqn (15). This approximation is reasonable for modelling the mobility of lanthanide tags; this was recently demonstrated by Shishmarev and Otting for the “two-hinged” approximation, where the orientation of the susceptibility tensor is the same for each rotamer.<sup>29</sup>

### 4.1 General solution

The most straightforward way to solve eqn (15) analytically is to expand the probability density in spherical harmonics:

$$\rho(\mathbf{r}) = \sum_{l,m} \alpha_l^m(r) Y_l^m(\hat{\mathbf{r}}) \quad (20)$$

where  $\alpha_l^m(r)$  are radial functions serving as expansion coefficients in this angular function series:

$$\alpha_l^m(r) = \int_{\Omega} Y_l^{m*}(\hat{\mathbf{r}}) \rho(\mathbf{r}) d^2\hat{\mathbf{r}} \quad (21)$$

For a spherically isotropic probability density distribution, the sum in eqn (20) only has one term with  $l = 0$  and  $m = 0$ ; for a spherically anisotropic density, higher rank terms would also be present.

The Fourier transform of the density in eqn (20) leaves the angular part the same but the radial part is integrated with spherical Bessel functions of the first kind:

$$\hat{\rho}(\mathbf{k}) = \sqrt{\frac{2}{\pi}} \sum_{l,m} (-i)^l Y_l^m(\hat{\mathbf{k}}) \int_0^\infty \alpha_l^m(s) j_l(ks) s^2 ds \quad (22)$$

We shall substitute this expression into eqn (16) and expand the products of spherical harmonics using Clebsch–Gordan coefficients. The result has three terms when  $l \geq 2$  and two when  $l < 2$ :

$$Y_l^m Y_2^{m'} = \sqrt{\frac{5(2l+1)}{4\pi}} \left( \frac{C_{l,0,2,0}^{l-2,0} C_{l,m,2,m'}^{l-2,m+m'}}{\sqrt{2l-3}} Y_{l-2}^{m+m'} + \frac{C_{l,0,l',0}^{l,0} C_{l,m,l',m'}^{l,m+m'}}{\sqrt{2l+1}} Y_l^{m+m'} + \frac{C_{l,0,l',0}^{l+2,0} C_{l,m,l',m'}^{l+2,m+m'}}{\sqrt{2l+5}} Y_{l+2}^{m+m'} \right) \quad (23)$$

With this substitution in place, eqn (16) acquires the following form:

$$\hat{\sigma}(\mathbf{k}) = -\frac{\sqrt{10}}{6\pi} \sum_{l,m} (-i)^l \sqrt{2l+1} \left[ \int_0^\infty j_l(ks) \alpha_l^m(s) s^2 ds \right] \times \sum_{m'=-2}^2 \chi_{m'} \left[ \frac{C_{l,0,2,0}^{l-2,0} C_{l,m,2,m'}^{l-2,m+m'}}{\sqrt{2l-3}} Y_{l-2}^{m+m'}(\hat{\mathbf{k}}) + \frac{C_{l,0,l',0}^{l,0} C_{l,m,l',m'}^{l,m+m'}}{\sqrt{2l+1}} Y_l^{m+m'}(\hat{\mathbf{k}}) + \frac{C_{l,0,2,0}^{l+2,0} C_{l,m,2,m'}^{l+2,m+m'}}{\sqrt{2l+5}} Y_{l+2}^{m+m'}(\hat{\mathbf{k}}) \right] \quad (24)$$



To obtain the final answer, we must take the inverse Fourier transform of eqn (24). The spherical harmonics again remain the same, and the radial part is integrated with spherical Bessel functions. The double integrals that make an appearance are all straightforward. Changing the order of integration and integrating the product of two spherical Bessel functions with respect to  $k$  yields:

$$\int_0^\infty \int_0^\infty \alpha_l^m(s) s^2 j_l(ks) j_{l-2}(kr) k^2 ds dk \quad (25)$$

$$= \pi \left( l - \frac{1}{2} \right) r^{l-2} \int_r^\infty \alpha_l^m(s) s^{-l+1} ds$$

$$\int_0^\infty \int_0^\infty \alpha_l^m(s) s^2 j_l(ks) j_l(kr) k^2 ds dk \quad (26)$$

$$= \frac{\pi}{2r^2} \int_0^\infty \alpha_l^m(s) s^2 \delta(r-s) ds = \frac{\pi}{2} \alpha_l^m(r)$$

$$\int_0^\infty \int_0^\infty \alpha_l^m(s) s^2 j_l(ks) j_{l+2}(kr) k^2 ds dk \quad (27)$$

$$= \pi \left( l + \frac{3}{2} \right) r^{-l-3} \int_0^r \alpha_l^m(s) s^{l+2} ds$$

Taking everything together we obtain the general analytical solution for eqn (15):

$$\sigma(\mathbf{r}) = -\frac{\sqrt{20}}{12\sqrt{\pi}} \sum_{l,m} \sqrt{2l+1} \quad (28)$$

$$\times \sum_{m'=-2}^2 \chi_{m'} \left( P_{l,m,m'}^{\text{OUT}}(\mathbf{r}) + P_{l,m,m'}^0(\mathbf{r}) + P_{l,m,m'}^{\text{IN}}(\mathbf{r}) \right)$$

There are three physically different contributions associated with the three integrals of the radial probability density functions in eqn (25)–(27). The first one corresponds to the integral of the density weighted with a monotonically decreasing function over the region outside the sphere of radius  $r$ :

$$P_{l,m,m'}^{\text{OUT}}(\mathbf{r}) = -\frac{2l-1}{\sqrt{2l-3}} C_{l,0,2,0}^{l-2,0} C_{l,m,2,m'}^{l-2,m+m'} Y_{l-2}^{m+m'}(\hat{\mathbf{r}}) r^{l-2} \int_r^\infty \alpha_l^m(s) s^{-l+1} ds \quad (29)$$

For any nucleus positioned outside the bounding sphere of the paramagnetic centre probability density this contribution is zero. The second contribution is proportional to the probability density at the nucleus itself:

$$P_{l,m,m'}^0(\mathbf{r}) = \frac{1}{\sqrt{2l+1}} C_{l,0,2,0}^{l,0} C_{l,m,2,m'}^{l,m+m'} Y_{l+2}^{m+m'}(\hat{\mathbf{r}}) \alpha_l^m(\mathbf{r}) \quad (30)$$

We are in practice unlikely to be able to measure chemical shifts of the nuclei that are directly underneath the spin density due to their fast relaxation. We are thus left with the third and the most important contribution that is associated with the part of the density that is inside the sphere of radius  $r$ :

$$P_{l,m,m'}^{\text{IN}}(\mathbf{r}) = -\frac{2l+3}{\sqrt{2l+5}} C_{l,0,2,0}^{l+2,0} C_{l,m,2,m'}^{l+2,m+m'} Y_{l+2}^{m+m'}(\hat{\mathbf{r}}) r^{-l-3} \int_0^r \alpha_l^m(s) s^{l+2} ds \quad (31)$$

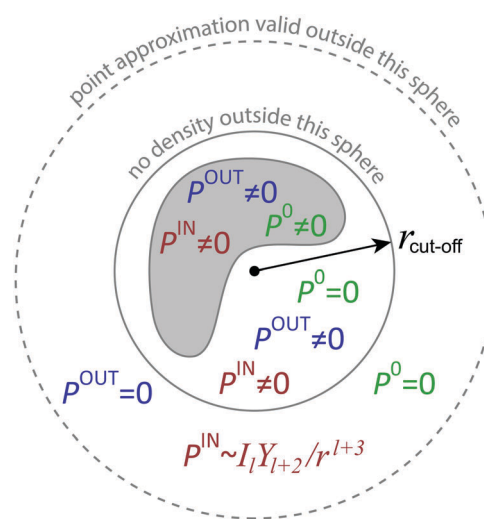


Fig. 1 A schematic representation of the paramagnetic centre probability density (grey shaded area) and regions where different terms in eqn (28) contribute to the resulting pseudocontact shift. Outside the bounding sphere (grey solid line) only  $P^{\text{IN}}$  is non-zero; inside the bounding sphere an additional term  $P^{\text{OUT}}$  becomes important;  $P^0$  contributes only inside the density. At a sufficient distance from the bounding sphere (grey dashed line) the point paramagnetic centre approximation becomes valid.

If we assume that there is no paramagnetic centre density outside that sphere, this would allow us to extend the upper integration limit to infinity and the integrals then correspond to the multipole moments of the probability density of the paramagnetic centre:

$$I_l^m = \int_0^\infty \left[ \int_\Omega Y_l^m(\hat{s}) \rho(\mathbf{s}) d^2 \hat{s} \right] s^{l+2} ds \quad (32)$$

A schematic diagram of the entire argument is given in Fig. 1. Outside the bounding sphere, the final expression for the PCS is:

$$\sigma(\mathbf{r}) = \frac{\sqrt{20}}{12\sqrt{\pi}} \sum_{l,m} \frac{1}{r^{l+3}} \frac{\sqrt{2l+1}(2l+3)}{\sqrt{2l+5}} I_l^m C_{l,0,2,0}^{l+2,0} \quad (33)$$

$$\times \sum_{m'=-2}^2 \chi_{m'} C_{l,m,2,m'}^{l+2,m+m'} Y_{l+2}^{m+m'}(\hat{\mathbf{r}})$$

This solution tells us that the PCS is sensitive to the multipole moments of the paramagnetic centre probability density. In the case of isotropic distribution, the PCS is the same as from the point source – this follows from eqn (33) when we put  $l$  to zero; it simplifies into eqn (7). Therefore, only the anisotropy of the density makes a difference in the PCS compared to the point source. We shall therefore proceed to explore some simple anisotropic probability densities.

A more detailed derivation of eqn (28) and (33) is provided in the ESI.†

## 4.2 PCS from a Gaussian paramagnetic centre distribution

This section contains an analysis of the consequences of simple deviations from the point paramagnetic centre approximation. To get a quantitative idea about how far the point dipole



approximation can in practice be stretched, we shall consider a family of Gaussian paramagnetic centre distributions.

It follows from the treatment above that the PCS field outside any isotropic paramagnetic centre distribution is identical to the PCS field generated by a point centre. The difference from the point PCS appears only inside the density. For example, in the case of an isotropic Gaussian, where the sum in eqn (20) has only one element with  $l = 0$  and  $m = 0$ , we have

$$\sigma_0^0(r) = \frac{1}{a^3\pi\sqrt{2}} e^{-\frac{r^2}{2a^2}} \quad (34)$$

and the exact PCS is:

$$\begin{aligned} \sigma(\mathbf{r}) = & \frac{\sqrt{2}}{24\pi\sqrt{\pi}r^3a^3} \left( 3\sqrt{2}\pi a^3 \operatorname{erf}\left(\frac{r}{\sqrt{2}a}\right) - 2re^{-\frac{r^2}{2a^2}}(r^2 + 3a^2) \right) \\ & \times \sum_{m=-2}^2 \chi_m Y_2^m(\hat{\mathbf{r}}) \end{aligned} \quad (35)$$

where  $a$  is the standard deviation of the Gaussian. Fig. 2 demonstrates that a significant difference between the point and the isotropic Gaussian case appears only for  $r < 3a$ , which is not particularly interesting or dangerous because nuclei at such short distances from the paramagnetic centre are not normally visible in PCS experiments due to their rapid transverse relaxation.

A rather more worrying situation emerges when we consider a realistically anisotropic Gaussian distribution of the paramagnetic centre:

$$\rho(x, y, z) = \frac{1}{a^2c\sqrt{(2\pi)^3}} \exp\left[-\frac{1}{2}\left(\frac{x^2 + y^2}{a^2} + \frac{z^2}{c^2}\right)\right] \quad (36)$$

The case with  $a > c$  corresponds to an oblate ellipsoid and the case with  $a < c$  to a prolate one. The anisotropy results in the emergence of the next multipole in eqn (32):

$$I_2^0 = \sqrt{5}(c^2 - a^2)/(2\sqrt{\pi}) \quad (37)$$

with the amplitude that depends on the difference in the squares of axial and equatorial sizes of the Gaussian. The resulting correction to the PCS

$$\sigma_2^0(\mathbf{r}) = \frac{\sqrt{5}(c^2 - a^2)}{24\pi} \frac{1}{r^5} \sum_{m=-2}^2 \chi_m \sqrt{(m^2 - 4)(m^2 - 3^2)} Y_4^m(\hat{\mathbf{r}}) \quad (38)$$

involves fourth rank spherical harmonics and the fifth power of the distance. This is illustrated graphically in Fig. 2 – it is clear that the correction to the point dipole solution is larger than 1 ppm up to 10 Å away from the spin label location.

The magnitudes and the asymptotic behaviour of the three contributions to the PCS in eqn (28) can be summarised by the following rules:

- (1) For any nucleus located outside the sphere that is three times the radius of the bounding sphere of the paramagnetic centre probability density, the point paramagnetic centre approximation is valid.
- (2) For any nucleus located in the immediate vicinity of the bounding sphere, extra multipoles would appear in the PCS, and eqn (33) must be used.
- (3) For any nucleus located inside the bounding sphere, no simplifications are available, and the full analytical solution in eqn (28) must be used.

## 5. Analytical solution to the inverse problem

Eqn (33) is linear with respect to the components of the susceptibility tensor and the multipole moments of the probability density. Both sets of parameters are therefore easy to extract by fitting a sufficiently large dataset comprising nuclear coordinates and pseudocontact shifts. However, because experimental PCS measurements inside the bounding sphere of the paramagnetic centre density are not usually realistic, only the

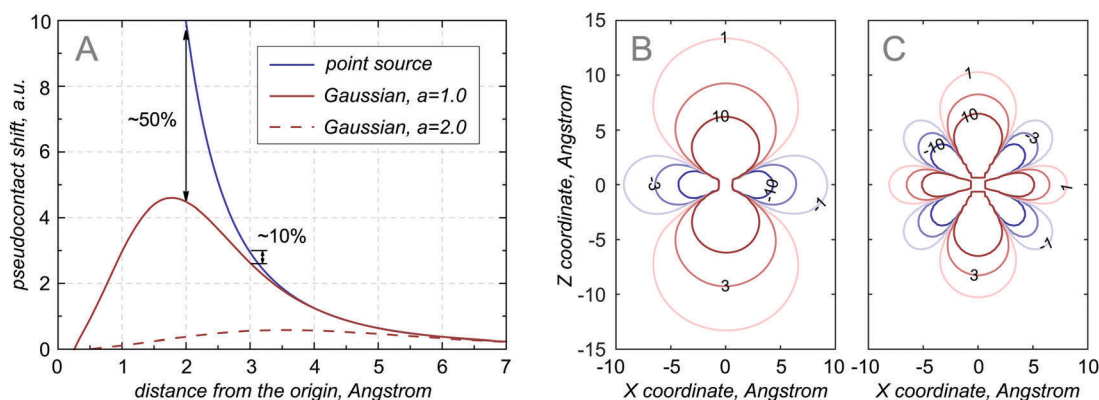
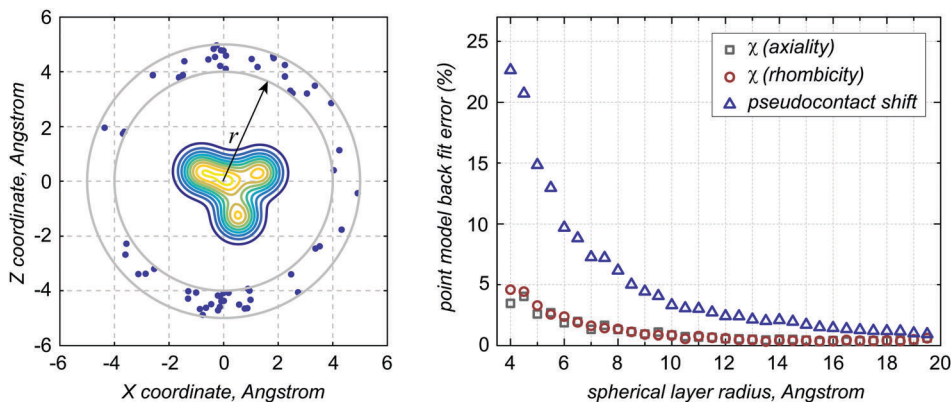


Fig. 2 (A) Amplitude of the radial part of the pseudocontact shift field as a function of distance from a point source (blue line) and from isotropic Gaussian sources with standard deviations of 1 Å (solid red line) and 2 Å (dashed red line). (B and C) PCS isolines in ppm from a prolate Gaussian source with  $a = 1$  Å,  $c = 3$  Å and  $\chi = \text{diag}([-0.01, -0.02, 0.03]) \text{Å}^3$ . The Z axis of the Gaussian is parallel to the Z axis of the susceptibility tensor eigenframe. The contribution to the PCS from the isotropic part of the density is on the left and the contribution arising from the anisotropic part is on the right.





**Fig. 3** Left panel: a schematic of the model system used to analyse the behaviour of the multipole terms in eqn (33) – the 2D slice in the XZ plane shows the contour lines of the paramagnetic centre distribution formed by four isotropic Gaussians with a standard deviation of  $0.5 \text{ \AA}$  scattered randomly within a  $3 \times 3 \times 3 \text{ \AA}$  cube; 300 nuclei (blue dots) are scattered randomly within a  $1 \text{ \AA}$  layer at varying distances from the origin. Right panel: point model fractional back fit error in the pseudocontact shifts and the susceptibility tensor as a function of the radius of the spherical layer in which the nuclei are scattered.

multipole moments (rather than the density itself) may be extracted in a well-defined way.

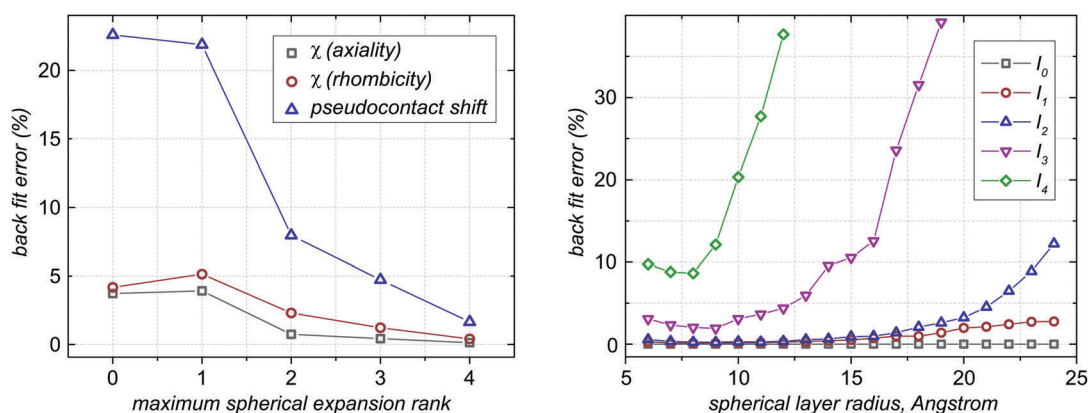
An illustration for 300 nuclei randomly placed around a lump of paramagnetic centre probability density is given in Fig. 3: the density was formed by four isotropic Gaussians with a standard deviation of  $0.5 \text{ \AA}$  placed randomly within a  $3 \times 3 \times 3 \text{ \AA}$  cube; nuclei were scattered within a  $1 \text{ \AA}$  thick spherical layer at varying distances from the origin; the axiality and the rhombicity of the susceptibility tensor were chosen randomly from the typical range reported in the literature:<sup>34</sup>  $\chi_{\text{ax}} = -0.45 \text{ \AA}^3$  and  $\chi_{\text{rh}} = -0.1 \text{ \AA}^3$ . PCS amplitudes in this system range from  $\pm 90 \text{ ppm}$  (at  $5 \text{ \AA}$  distance) to  $\pm 1 \text{ ppm}$  (at  $20 \text{ \AA}$  distance).

Pseudocontact shifts in this model system were calculated exactly and then an attempt was made to back-fit eqn (33) truncated at different spherical ranks. The results were averaged over 50 instances of random nuclear position sets. As expected, the point model does not perform well below  $10 \text{ \AA}$  from the bounding sphere (Fig. 3, right panel). Still, even though the PCS values are badly reproduced, the susceptibility tensor is quite

resilient – even at the distance of  $4 \text{ \AA}$  it is recovered to within  $\sim 5\%$  accuracy.

It is clear from Fig. 3 that, at distances comparable to the size of the paramagnetic centre probability distribution, the point model breaks down and further terms are required in eqn (33). Their beneficial effect is illustrated in Fig. 4 – adding terms of higher spherical rank to the expansion dramatically reduces the PCS back-fit error in the vicinity of the bounding sphere.

Due to the steep distance dependence and the wobbly angle dependence of the higher multipole terms in eqn (33), the accuracy with which these terms may be extracted also falls steeply as the nuclei are moved further away from the bounding sphere of the paramagnetic centre probability density. This is illustrated in Fig. 4 – it is clear that multipole moments of spherical rank higher than 4 cannot be reliably extracted no matter how close the nuclei are to the bounding sphere. We would not therefore recommend taking eqn (33) beyond the fourth spherical rank.



**Fig. 4** Left panel: spherical rank dependence of the fractional back fit error produced by eqn (33) in the case when the nuclei are scattered in the spherical layer between  $4 \text{ \AA}$  and  $5 \text{ \AA}$  away from the origin (a schematic is given in Fig. 3). Right panel: fractional back fit error in the multipole moments of the paramagnetic centre probability density distribution in eqn (33) as a function of the radius of the spherical layer containing the nuclei (a schematic is given in Fig. 3).



## 6. Numerical solution to the direct problem

If the probability density of the paramagnetic centre is discretised on a finite three-dimensional grid, two general numerical avenues become available for the solution of eqn (15) – the finite-difference method and the Fourier transform method. The latter has the advantage of being fast, but the disadvantage of requiring periodic boundary conditions. This section explores both methods and comments on their relative merits.

### 6.1 Finite difference methods

A general algorithm for the generation of elementary finite difference operators of a given derivative and accuracy order on a given finite grid has been published by Fornberg.<sup>35</sup> For a rectangular grid with  $N$  points in the X direction,  $M$  points in the Y direction and  $K$  points in the Z direction, finite difference matrix representations of the relevant derivative operators acting on the vectorisation of the density cube are:

$$\begin{aligned} \left[ \frac{\partial}{\partial x} \right] &= \mathbf{D}_N^{(1)} \otimes \mathbf{1}_M \otimes \mathbf{1}_K, & \left[ \frac{\partial^2}{\partial x^2} \right] &= \mathbf{D}_N^{(2)} \otimes \mathbf{1}_M \otimes \mathbf{1}_K, \\ \left[ \frac{\partial^2}{\partial x \partial y} \right] &= \mathbf{D}_N^{(1)} \otimes \mathbf{D}_M^{(1)} \otimes \mathbf{1}_K \end{aligned} \quad (39)$$

and similarly for the other first and second derivatives. In these expressions,  $\mathbf{D}_N^{(k)}$  is a matrix representation of the  $k$ -th derivative operator on a grid with  $N$  points, and  $\mathbf{1}_M$  is a unit matrix of dimension  $M$ . The dimension of the matrices in eqn (39) is  $NMK$ . For a typical grid with 256 points in each dimension this is a large number, but because finite difference operators are local, the matrices produced by eqn (39) are very sparse.

The matrices required for the solution of eqn (15) are  $\mathbf{K} = -[\vec{\nabla}^T \cdot \chi \cdot \vec{\nabla}] / 3$  and  $\mathbf{L} = [\vec{\nabla}^T \cdot \vec{\nabla}]$ . In terms of these matrices, the solution may be written as

$$\sigma = \mathbf{L}^{-1} \mathbf{K} \rho \quad (40)$$

where  $\rho$  is the vectorization of the probability density cube on the chosen grid and  $\sigma$  is the vectorisation of the pseudocontact shift field on the same grid; here and below we take only the traceless part of the  $\chi$  tensor. The final step is to project out the PCS values on the nuclei, for which the interpolation matrix  $\mathbf{P}_N$ , also published by Fornberg in the same paper,<sup>35</sup> may be used. The final expression is

$$\sigma_N = \mathbf{P}_N \mathbf{L}^{-1} \mathbf{K} \rho \quad (41)$$

where the action by the very sparse inverse Laplacian matrix on the  $\mathbf{K} \rho$  vector is best computed using an iterative solver, such as GMRES.<sup>36</sup> Here and below the  $\mathbf{L}^{-1}$  symbol should be understood in that sense – the inverse Laplacian is never computed explicitly.

This method is very easy to set up – see, for example, the *kpcs.m* function supplied with *Spinach* library.<sup>37</sup> Its downside is unfavourable scaling: the dimension of the finite difference matrices involved is  $KMN$  and the scaling of sparse solvers is

approximately quadratic in the matrix dimension, meaning that the numerical complexity of this method scales approximately as the sixth power of the grid size. On a contemporary computer, a solution for a  $256 \times 256 \times 256$  point grid takes about an hour.

### 6.2 Fourier transform methods

An alternative method for solving eqn (15) takes advantage of the  $k$ -space expression in eqn (16), where the Fourier transform can be taken numerically using the fast Fourier transform (FFT) algorithm.<sup>38</sup> Using the same interpolation operator as in eqn (41), we get

$$\sigma_N = -\frac{1}{3} \mathbf{P}_N \cdot \text{Re} \left[ \text{FFT}_- \left\{ \frac{\mathbf{k}^T \cdot \chi \cdot \mathbf{k}}{\mathbf{k}^T \cdot \mathbf{k}} \text{FFT}_+ \{ \rho(\mathbf{r}) \} \right\} \right] \quad (42)$$

where  $\chi$  stands for the traceless part of the magnetic susceptibility tensor. This method is about the same as the finite difference method in terms of the implementation complexity, but it is much faster – the complexity scaling of the three-dimensional fast Fourier transform is  $NMK \log(N) \log(M) \log(K)$ , which is close to cubic scaling with respect to the grid size. This takes the simulation time into the region of seconds. This method is particularly fast on modern computing hardware because fast Fourier transform routines are available for GPGPU coprocessor cards (*Nvidia Tesla K40* cards were used in this work). The performance gain is shown in the left panel of Fig. 5.

The only significant problem with the FFT method is the periodic boundary condition – care must be taken to ensure that the images do not contribute significantly to the solution. Thankfully, the PCS decays cubically with distance; it is sufficient for the distance from the paramagnetic centre to the cube boundary and the distance from the nucleus to the cube boundary to be about three times larger than the distance from the paramagnetic centre to the nucleus – a numerical example is given in Fig. 5.

## 7. Numerical solution to the inverse problem

The “inverse problem” refers to the task of reconstructing the paramagnetic centre probability density from the experimental values of pseudocontact shifts and atomic coordinates. The task may be formulated as finding the probability density  $\rho(\mathbf{r})$  that minimises the following functional:

$$\Omega[\rho(\mathbf{r})] = E[\rho(\mathbf{r})] + T[\rho(\mathbf{r})], \quad (43)$$

where the least squares error is

$$E[\rho(\mathbf{r})] = \|\mathbf{P}_N \mathbf{L}^{-1} \mathbf{K} \rho(\mathbf{r}) - \sigma_{\text{expt}}\|^2 \quad (44)$$

and the standard Tikhonov regularisation term emphasizing smooth solutions is<sup>39</sup>

$$T[\rho(\mathbf{r})] = \lambda \|\mathbf{L} \rho(\mathbf{r})\|^2, \quad (45)$$

where  $\lambda$  is a user-specified regularisation parameter. Regularisation is only needed when the inverse problem is ill-posed, *i.e.* the number of points in the probability density grid exceeds



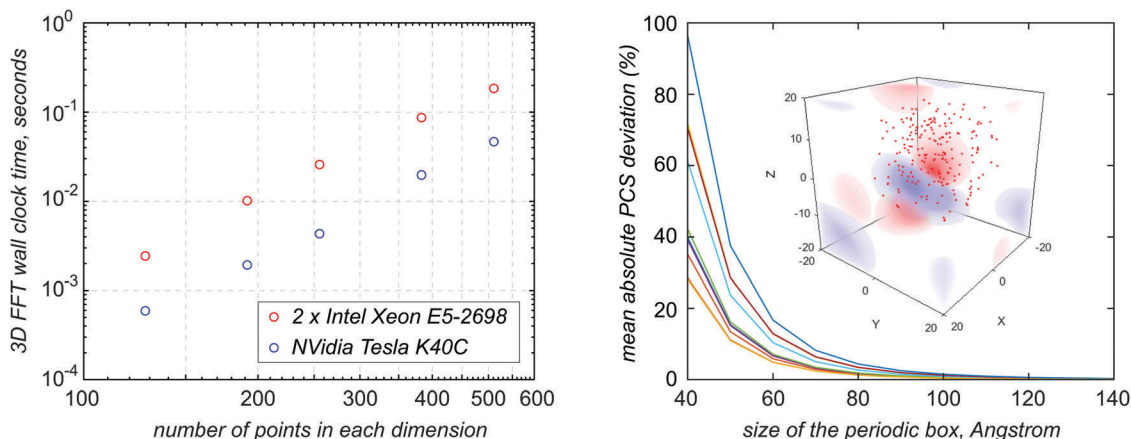


Fig. 5 Left panel: 3D FFT wall clock time as a function of the grid size for the calculation performed on the CPU vs. a dedicated coprocessor card. Right panel: PCS error due to the periodic boundary conditions in the FFT method as a function of the periodic box size. The nuclei (red points in the inset) are placed randomly in the 20 Å box in a positive octant; the paramagnetic centre has an isotropic Gaussian distribution centred at  $[-10, -10, -10]$  Å with a standard deviation of 1.0 Å.

the number of experimental PCS points and/or all experimental points are outside the bounding sphere of the density. In that case the recovered density is only an approximation that agrees with the experimental data and has the required smoothness at the same time – this point is discussed further in Section 8.

### 7.1 Finite difference method

In order to proceed with minimising  $\Omega[\rho(\mathbf{r})]$  with respect to  $\rho(\mathbf{r})$ , the state-of-the-art Newton–Raphson minimisers require analytical expressions for their first and second variation.<sup>40</sup> In a matrix representation on a discrete grid, these variations become the gradient vector and the Hessian matrix. The following vector relations are useful as a starting point.

$$\begin{aligned} \frac{\partial}{\partial \mathbf{x}^T} \|\mathbf{Ax} - \mathbf{b}\|^2 &= 2\mathbf{A}^T(\mathbf{Ax} - \mathbf{b}) \\ \frac{\partial^2}{\partial \mathbf{x}^T \partial \mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|^2 &= 2\mathbf{A}^T \mathbf{A} \end{aligned} \quad (46)$$

If we use  $\boldsymbol{\rho}$  to denote the vectorised form of a discrete representation of  $\rho(\mathbf{r})$  on a finite three-dimensional grid, the expressions for the gradient and the Hessian of the least squares error part of eqn (43) are:

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\rho}^T} \|\mathbf{PL}^{-1}\mathbf{K}\boldsymbol{\rho} - \boldsymbol{\sigma}_{\text{expt}}\|^2 &= 2\mathbf{KL}^{-1}\mathbf{P}^T(\mathbf{PL}^{-1}\mathbf{K}\boldsymbol{\rho} - \boldsymbol{\sigma}_{\text{expt}}) \\ &= 2\mathbf{KL}^{-1}\mathbf{P}^T(\boldsymbol{\sigma}_{\text{theo}} - \boldsymbol{\sigma}_{\text{expt}}) \\ \frac{\partial^2}{\partial \boldsymbol{\rho}^T \partial \boldsymbol{\rho}} \|\mathbf{PL}^{-1}\mathbf{K}\boldsymbol{\rho} - \boldsymbol{\sigma}_{\text{expt}}\|^2 &= 2\mathbf{KL}^{-1}\mathbf{P}^T\mathbf{PL}^{-1}\mathbf{K} \end{aligned} \quad (47)$$

where advantage was taken of the fact that  $\mathbf{L}$  and  $\mathbf{K}$  are symmetric matrices and that  $\boldsymbol{\sigma}_{\text{theo}} = \mathbf{PL}^{-1}\mathbf{K}\boldsymbol{\rho}$  would normally already be computed and stored in memory by the time the derivatives are requested. Although the Hessian in eqn (47) is formally a matrix of a very large dimension, in practice only the result of its multiplication by a vector is required by the modern Newton–Raphson optimisers because they employ iterative sparse linear solvers for the inverse-Hessian-times-gradient operation.<sup>41,42</sup>

Similarly, the gradient and the Hessian for the Tikhonov regularisation term are:

$$\frac{\partial}{\partial \boldsymbol{\rho}^T} T[\boldsymbol{\rho}] = 2\lambda \mathbf{L}^T \mathbf{L} \boldsymbol{\rho}, \quad \frac{\partial^2}{\partial \boldsymbol{\rho}^T \partial \boldsymbol{\rho}} T[\boldsymbol{\rho}] = 2\lambda \mathbf{L}^T \mathbf{L} \quad (48)$$

It is in practice unnecessary to consider every point of the grid to be a variable. The location of the paramagnetic centre is usually known approximately, and it is sufficient to only consider the points in the immediate vicinity – typically a  $20 \times 20 \times 20$  Å cube around the expected location.

### 7.2 Fourier transform method

About a factor of a hundred in performance may be gained (at the cost of worrying about the periodic boundary as described in Section 6.2) by using the Fourier solution in eqn (42). The overall error functional is still the same as in eqn (43), but the terms for the least squares error and the Tikhonov regularisation now involve three-dimensional fast Fourier transforms:

$$\begin{aligned} E[\rho(\mathbf{r})] &= \|\boldsymbol{\sigma}_{\text{theo}} - \boldsymbol{\sigma}_{\text{expt}}\|^2, \\ T[\rho(\mathbf{r})] &= \lambda \text{Re}[\text{FFT}_- \{ \mathbf{k} \cdot \mathbf{k}^T \text{FFT}_+ \{ \rho(\mathbf{r}) \} \}]^2 \end{aligned} \quad (49)$$

The first variation and the action by the second variation of  $E[\rho(\mathbf{r})]$  on a probe function  $\eta(\mathbf{r})$  are:

$$\begin{aligned} \frac{\delta E}{\delta \rho} &= -\frac{2}{3} \text{Re} \left[ \text{FFT}_- \left\{ \frac{\mathbf{k}^T \cdot \boldsymbol{\chi} \cdot \mathbf{k}}{\mathbf{k} \cdot \mathbf{k}^T} \text{FFT}_+ \{ \mathbf{P}_N^T [\boldsymbol{\sigma}_{\text{theo}} - \boldsymbol{\sigma}_{\text{expt}}] \} \right\} \right] \\ \frac{\delta^2 E}{\delta \rho^2} [\eta(\mathbf{r})] &= -\frac{2}{3} \text{Re} \left[ \text{FFT}_- \left\{ \frac{\mathbf{k}^T \cdot \boldsymbol{\chi} \cdot \mathbf{k}}{\mathbf{k} \cdot \mathbf{k}^T} \text{FFT}_+ \left\{ \mathbf{P}_N^T \mathbf{P}_N \text{FFT}_- \left\{ \frac{\mathbf{k}^T \cdot \boldsymbol{\chi} \cdot \mathbf{k}}{\mathbf{k} \cdot \mathbf{k}^T} \text{FFT}_+ \{ \eta(\mathbf{r}) \} \right\} \right\} \right\} \right] \end{aligned} \quad (50)$$



The corresponding variations for the Tikhonov regularisation functional are:

$$\frac{\delta T}{\delta \rho} = 2\lambda \text{Re} \left[ \text{FFT}_- \left\{ (\mathbf{k} \cdot \mathbf{k}^T)^2 \text{FFT}_+ \{ \rho(\mathbf{r}) \} \right\} \right] \quad (51)$$

$$\frac{\delta^2 T}{\delta \rho^2} [\eta(\mathbf{r})] = 2\lambda \text{Re} \left[ \text{FFT}_- \left\{ (\mathbf{k} \cdot \mathbf{k}^T)^2 \text{FFT}_+ \{ \eta(\mathbf{r}) \} \right\} \right]$$

A numerical technicality in eqn (49) and (50) concerns the behaviour of the  $\mathbf{k}^T \cdot \boldsymbol{\chi} \cdot \mathbf{k} / \mathbf{k} \cdot \mathbf{k}^T$  term around the origin of the  $k$ -space. It is easy to demonstrate that the limit value

$$\lim_{|\mathbf{k}| \rightarrow 0} \frac{\mathbf{k}^T \cdot \boldsymbol{\chi} \cdot \mathbf{k}}{\mathbf{k} \cdot \mathbf{k}^T} \quad (52)$$

corresponds to the integral of the solution in the real space. In the case of the pseudocontact shift this integral is known to be zero; this resolves the ambiguity.

### 7.3 Regularisation parameters

The standard way of choosing the Tikhonov regularisation parameter is the  $L$ -curve method:<sup>43,44</sup> the error functional minimisation is repeated with different values of  $\lambda$  and the value is picked that corresponds to the maximum curvature  $\kappa(\lambda)$  of the following curve

$$\begin{cases} x(\lambda) = \ln(E_i[\rho_{\text{opt}}]) \\ y(\lambda) = \ln(T_i[\rho_{\text{opt}}]) \end{cases} \quad \kappa(\lambda) = \frac{|x'y'' - x''y'|}{(x'^2 + y'^2)^{3/2}} \quad (53)$$

In practice, a discrete point set  $\{E_{\lambda_k}, T_{\lambda_k}\}$  for different regularisation parameter values  $\lambda_k$  is interpolated by a fifth order spline and then differentiated numerically. The standard third order spline is not sufficient here because second derivatives are involved in the definition of the curvature that later needs to be differentiated again to find a maximum. The derivatives of the spline are fed into the expression for the curvature and the maximum is found with respect to  $\lambda$  using standard optimisation techniques.

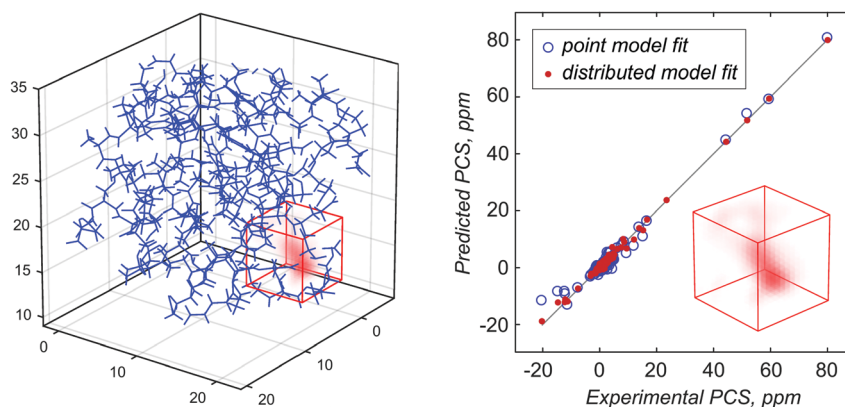
## 8. Probability density reconstruction examples

Paramagnetic centre probability density reconstruction and its interpretation for real-life paramagnetic protein systems is a very large block of work that has been submitted for publication as a separate paper. This paper deals with the technical side of the matter; here we shall therefore only look at two basic example cases and make some general cautionary comments.

An example, shown in Fig. 6, deals with the paramagnetic centre density distribution in the well characterized calbindin  $D_{9K}$  in which one or both of the two calcium ions have been replaced by a thulium ion.<sup>45</sup> Detailed coordinate and PCS data for this protein are available in the ESI of the paper by the Florence group.<sup>34</sup> The point paramagnetic centre model produces a rather unsatisfactory fit in the vicinity of the metal (Fig. 6, right panel). A reasonable explanation (proposed to the authors by Gottfried Otting and Thomas Huber) would be that fast exchange between metal binding sites could be present. We can now confirm that there is indeed an elongated distribution in the metal position (Fig. 6, left panel).

Simple and intuitive though the reconstructions of the kind shown in Fig. 6 might appear, a few words of caution are in order. An important consequence of the asymptotic behaviour of the three components of the exact solution in eqn (28) is that the fine internal details of the paramagnetic centre probability density distribution are expected to be lost if the PCS data are measured for the nuclei that are positioned outside the bounding sphere of the density. In practice this means that the shape of the bounding surface of the probability distribution is well reproduced by the reconstruction, but its internal details are not.

The consequences are illustrated in Fig. 7. The left diagram shows a simple test density composed of three Gaussian functions. The middle diagram shows the reconstruction performed using PCS data from the nuclei that are distributed randomly within a 10 Å cube around the origin, including inside the



**Fig. 6** Paramagnetic centre probability density reconstruction for  $\text{Tm}^{3+}$  labelled calbindin  $D_{9K}$  (PDB code: 1IGV),<sup>34,45</sup> for which the point approximation is known to produce a poor fit (right panel, blue circles). Numerical reconstruction (left panel) reveals that the thulium ion probability density has an elongated shape, likely as a consequence of local conformational mobility or multiple competing oxygen coordination sites. The red cube is the volume in which the probability density is allowed to vary during the Tikhonov regularised reconstruction process.



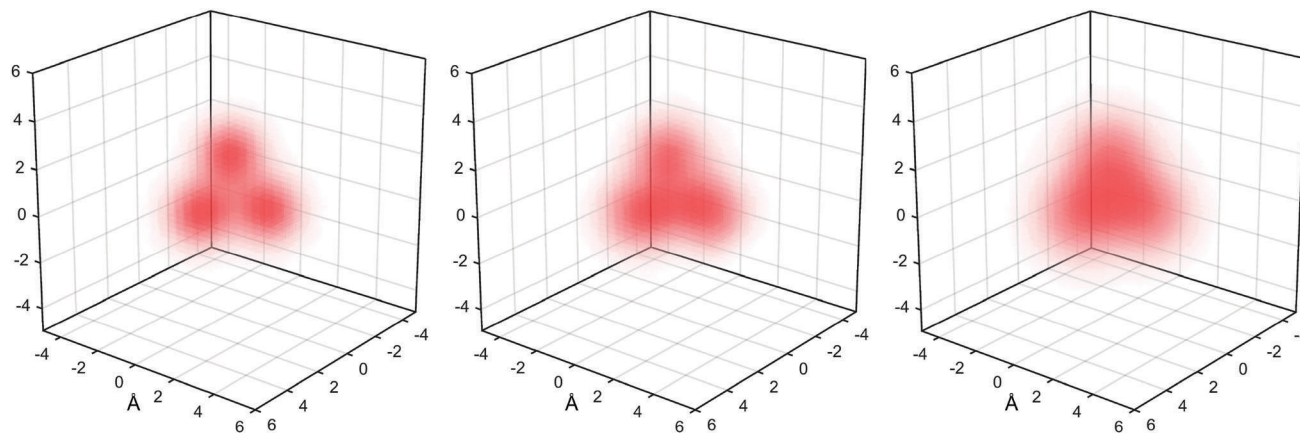


Fig. 7 Left panel: A trimodal paramagnetic centre probability density distribution composed of three isotropic Gaussians with a standard deviation of 1 Å placed at [2, 0, 0], [0, 2, 0] and [0, 0, 2] Angstroms. Middle panel: probability density reconstruction from the PCS values computed at 100 nuclei placed randomly throughout a  $10 \times 10 \times 10$  Å cube centred at the origin. Right panel: Probability density reconstruction from the PCS computed at 500 nuclei placed inside the same  $10 \times 10 \times 10$  Å cube, but outside the 5 Å sphere centred at density.

density. The right diagram shows the reconstruction performed using PCS data from the nuclei that are placed outside the 5 Å sphere around the density. The latter case is what typically happens in paramagnetic NMR; it is clear that a reasonable likeness of the true probability density is recovered, but the internal details of the distribution are lost because PCS values at those points are not sensitive to the radial functions in the probability density expansion in eqn (20). This is both the consequence of the structure of the multipolar expansion and the price to pay for a Tikhonov solution to what is in general an ill-posed problem.

## 9. Conclusions

Delocalisation of the paramagnetic centre creates deviations from the point dipole model for the pseudocontact shift. The analytical solution to the partial differential equation for PCS indicates that the key factor in such deviations is the angular anisotropy in the probability density – the PCS from spherical paramagnetic centre distributions is identical to the point dipole PCS outside the bounding sphere of the density.

Multipoles with spherical rank  $l$  in the paramagnetic centre distribution create rank  $l + 2$  multipoles in the resulting PCS field that decay with the distance as  $1/r^{l+3}$  outside the bounding sphere of the density. Due to fast decay of these higher rank components, only terms up to spherical rank four can in practice be extracted by fitting experimental data. The behaviour of the PCS field inside the bounding sphere is more complicated and may be analysed numerically; however, experimental measurements in such close proximity to the paramagnetic centre are rarely possible.

Discretising the paramagnetic centre probability density and the PCS field on finite grids makes it possible to solve the partial differential equation for PCS numerically, using either finite difference operators or fast Fourier transforms. The ill-posed inverse problem of reconstructing the probability density from experimental PCS data can then be solved using Tikhonov

regularisation under the assumption that the paramagnetic centre has the same effective magnetic susceptibility anisotropy at every point in the distribution.

The methods described in this paper are implemented in versions 1.8 and later of the *Spinach* library.<sup>37</sup>

## Acknowledgements

This work was made possible by a grant from EPSRC (EP/N006895/1). Stimulating discussions with Thomas Huber, Alexander Karabanov, Claudio Luchinat, Gottfried Otting and David Parker are gratefully acknowledged.

## References

- 1 G. N. L. Mar, W. D. Horrocks and R. H. Holm, *NMR of Paramagnetic Molecules: Principles and Applications*, Elsevier Science, 1973.
- 2 I. Bertini, C. Luchinat and G. Parigi, *Solution NMR of Paramagnetic Molecules: Applications to Metallobiomolecules and Models*, Elsevier Science, 2001.
- 3 L. Banci, I. Bertini, G. G. Savellini, A. Romagnoli, P. Turano, M. A. Cremonini, C. Luchinat and H. B. Gray, *Proteins*, 1997, **29**, 68.
- 4 L. Banci, I. Bertini, L. K. Bren, A. M. Cremonini, B. H. Gray, C. Luchinat and P. Turano, *JBC, J. Biol. Inorg. Chem.*, 1996, **1**, 117.
- 5 L. Banci, I. Bertini, M. A. Cremonini, G. Gori-Savellini, C. Luchinat, K. Wüthrich and P. Güntert, *J. Biomol. NMR*, 1998, **12**, 553.
- 6 I. Bertini, C. Luchinat and G. Parigi, *NMR of Biomolecules: Towards Mechanistic Systems Biology*, 2012, p. 154.
- 7 M. A. S. Hass and M. Ubbink, *Curr. Opin. Struct. Biol.*, 2014, **24**, 45.
- 8 G. Otting, *Annu. Rev. Biophys.*, 2010, **39**, 387.



- 9 I. Bertini, Y. K. Gupta, C. Luchinat, G. Parigi, M. Peana, L. Sgheri and J. Yuan, *J. Am. Chem. Soc.*, 2007, **129**, 12786.
- 10 A. M. Kenwright, I. Kuprov, E. De Luca, D. Parker, S. U. Pandya, P. K. Senanayake and D. G. Smith, *Chem. Commun.*, 2008, 2514.
- 11 P. Harvey, K. H. Chalmers, E. De Luca, A. Mishra and D. Parker, *Chem. – Eur. J.*, 2012, **18**, 8748.
- 12 N. F. Ramsey, *Phys. Rev.*, 1950, **78**, 699.
- 13 N. F. Ramsey, *Phys. Rev.*, 1952, **86**, 243.
- 14 H. M. McConnell, *J. Chem. Phys.*, 1957, **27**, 226.
- 15 H. M. McConnell and R. E. Robertson, *J. Chem. Phys.*, 1958, **29**, 1361.
- 16 B. Bleaney, *J. Magn. Reson.*, 1972, **8**, 91.
- 17 R. M. Golding, R. O. Pascual and L. C. Stubbs, *Mol. Phys.*, 1976, **31**, 1933.
- 18 P. J. Stiles, *Mol. Phys.*, 1974, **27**, 501.
- 19 P. T. Stiles and R. M. Wing, *J. Magn. Reson.*, 1974, **15**, 510.
- 20 P. J. Stiles, *Proc. R. Soc. London, Ser. A*, 1975, **346**, 209.
- 21 P. J. Stiles, *Chem. Phys. Lett.*, 1975, **30**, 259.
- 22 F. Gendron, K. Sharkas and J. Autschbach, *J. Phys. Chem. Lett.*, 2015, **6**, 2183.
- 23 M. Kaupp, M. Bühl and V. G. Malkin, *Calculation of NMR and EPR Parameters: Theory and Applications*, Wiley, 2004.
- 24 L. F. Chibotaru and L. Ungur, *J. Chem. Phys.*, 2012, **137**, 064112.
- 25 D. Ganyushin and F. Neese, *J. Chem. Phys.*, 2013, **138**, 104113.
- 26 I. Bertini, C. Luchinat and G. Parigi, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2002, **40**, 249.
- 27 I. Bertini, P. Turano and A. J. Vila, *Chem. Rev.*, 1993, **93**, 2833.
- 28 C. Schmitz, R. Vernon, G. Otting, D. Baker and T. Huber, *J. Mol. Biol.*, 2012, **416**, 668.
- 29 D. Shishmarev and G. Otting, *J. Biomol. NMR*, 2013, **56**, 203.
- 30 G. T. P. Charnock and I. Kuprov, *Phys. Chem. Chem. Phys.*, 2014, **16**, 20184.
- 31 W. Van den Heuvel and A. Soncini, *J. Chem. Phys.*, 2013, **138**, 054113.
- 32 K. Sharkas, B. Pritchard and J. Autschbach, *J. Chem. Theory Comput.*, 2015, **11**, 538.
- 33 T. O. Pennanen and J. Vaara, *J. Chem. Phys.*, 2005, **123**, 174102.
- 34 I. Bertini, M. B. L. Janik, Y.-M. Lee, C. Luchinat and A. Rosato, *J. Am. Chem. Soc.*, 2001, **123**, 4181.
- 35 B. Fornberg, *Math. Comput.*, 1988, **51**, 699.
- 36 Y. Saad and M. H. Schultz, *SIAM J. Sci. Stat. Comput.*, 1986, **7**, 856.
- 37 H. J. Hogben, M. Krzystyniak, G. T. P. Charnock, P. J. Hore and I. Kuprov, *J. Magn. Reson.*, 2011, **208**, 179.
- 38 J. W. Cooley and J. W. Tukey, *Math. Comput.*, 1965, **19**, 297.
- 39 A. N. Tikhonov, A. Goncharsky, V. V. Stepanov and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Springer, Netherlands, 1995.
- 40 J. Nocedal and S. Wright, *Numerical optimization*, Springer Science & Business Media, 2006.
- 41 T. F. Coleman and Y. Li, *Math. Program.*, 1994, **67**, 189.
- 42 T. F. Coleman and Y. Li, *SIAM J. Optimiz.*, 1996, **6**, 418.
- 43 P. C. Hansen, *SIAM Rev.*, 1992, **34**, 561.
- 44 P. C. Hansen and D. P. O'Leary, *SIAM J. Sci. Comput.*, 1993, **14**, 1487.
- 45 S. Balayssac, B. Jiménez and M. Piccioli, *J. Biomol. NMR*, 2006, **34**, 63.

