9.3: Conversion of dipolar evolution data to distance distributions

Expression for the DEER signal

In Section 9.1.1 we have seen that the echo is modulated with $\cos(dt)$. Usually, this applies only to a fraction λ of the echo, because the pump pulse excites only a fraction λ of all spin packets of the coupling partner of the observer spin. Therefore, the echo signal for an isolated pair of electron spins in a fixed orientation θ with respect to the magnetic field is described by

$$F(t,r,\theta) = F(0)\{1-\lambda(\theta)[1-\cos(2d(r,\theta)t)]\} \tag{9.3.1}$$

where the dependence $d(\theta)$ is given by Eqs. (5.16) and (5.15). The dependence $\lambda(\theta)$ cannot be expressed in closed form, but often λ is so weakly correlated with θ that it can be assumed as a constant, empirical parameter. In this situation, Eq. (9.1) can be integrated over all orientations

$$F(t,r) = \int_0^{\pi/2} F(t,r,\theta)\sin\theta \mathrm{d}\theta \tag{9.3.2}$$

The pump pulse inverts not only the coupling partner of the observer spin in the same molecule, but also electron spins in remote other molecules. If these neighboring spins are homogeneously distributed in space, the background factor $B(t)$ that arises from them assumes the form

$$B(t) = \exp\left(-\frac{2\pi g^2 \mu_\mathrm{B}^2 \mu_0 N_\mathrm{A}}{9\sqrt{3}\hbar}\lambda' ct\right) \tag{9.3.3}$$

where the orientation-averaged inversion efficiency λ' is the fraction of spins excited by the pump pulse, g is an average g value, and c is the total concentration of spins. For subtle reasons, λ' differs significantly from the empirical two-spin modulation depth λ. Homogeneous distributions of neighboring spins that are nearly confined to a plane or a line give rise to a stretched exponential background function $B(t) = \exp\left[-(kt)^{D/3}\right]$, where D is a fractional dimension of the distribution that is usually somewhat larger than 2 or 1 for nearly planar or linear distributions, respectively. The total DEER signal is given by

$$V(t,r) = F(t,r)B(t) \tag{9.3.4}$$

If distance r is distributed with normalized probability density $P(r)$ $\left(\int_0^\infty P(r)\mathrm{d}r = 1\right)$, the form factor $F(t)$ needs to be replaced by $F_P(t) = \int_0^\infty P(r)F(t,r)\mathrm{d}r$.

Copy and Paste
Image here.
Delete this
placeholder image

Figure 9.3: Background correction in DEER spectroscopy. (a) Primary data $V(t)$ (simulation) normalized to $V(0)$. Dipolar modulation decays until a time t_dec. An exponentially decay function (red) is fitted to the data in the range $t_\mathrm{dec} \le t \le t_\mathrm{max}$, where $t_\mathrm{max} < \tau_2$ is the maximum dipolar evolution time. This background function $b(t)$ is extrapolated to the range $0 \le t < t_\mathrm{dec}$ (ochre). (b) The form factor $F(t)$ is obtained by normalizing the background function, $B(t) = b(t)/b(0)$ and dividing the normalized primary data $V(t)/V(0)$ by $B(t)$. It decays to a constant level $1-\lambda$, where λ is the modulation depth. The red curve is a simulation corresponding to the distance distribution extracted by Tikhonov regularization with optimum regularization parameter α.

Background correction

The information on the distance distribution $P(r)$ is contained in $F(t)$, which must thus be separated from $B(t)$. Often, the distribution is sufficiently broad for dipolar oscillations to decay within a time t_dec shorter than the maximum dipolar evolution time t_max (Fig. 9.3(a)). For $t_\mathrm{dec} \le t \le t_\mathrm{max}$, the primary signal is then given by $b(t) = (1-\lambda)\exp\left[-(kt)^{D/3}\right]$ plus noise. The expression for $b(t)$ is fitted to the primary data in this range (red line in Fig. 9.3(a)). In some cases, for instance for soluble proteins, a homogeneous distribution of molecules in three dimensions can be assumed, so that $D = 3$ can be fixed. Otherwise, D is treated as a fit parameter, as are k and λ. The background function $B(t)$ is obtained by extrapolating $b(t)$ to the range $0 \le t \le t_\mathrm{dec}$ (ochre line) and dividing it by $b(0) = 1-\lambda$. According to Eq. (9.4), the form factor $F(t)/F(0)$ results by dividing

$V(t)/V(0)$ by $B(t)$. For narrow distance distributions, oscillations in $V(t)/V(0)$ may endure until the longest attainable $t_{\max}$. This does not create a problem if at least the first oscillation is completed well before $t_{\max}$. All the following oscillations have very similar amplitude and do not bias the background fit. As a rule of thumb, a good estimate for $B(t)$ can be obtained by fitting data at $t \geq t_{\max}/2$ if $dt_{\max} \geq 4\pi$, i.e., if two full oscillations can be observed. If the data trace is shorter than that, background fitting is fraught with uncertainty. Wrong background correction may suppress long distances or create artificial peaks at long distances.

Tikhonov regularization with non-negativity constraint

In order to extract the distance distribution $P(r)$ from the experimental form factor $F(t)/F(0)$, we need to remove the constant contribution and renormalize to the dipolar evolution function

$$D(t) = \frac{F(t)/F(0) - (1-\lambda)}{\lambda} \tag{9.3.5}$$

and invert the integral equation $D(t) = \int_0^\infty P(r)K(t,r)\mathrm{d}r$, where the kernel $K(t,r)$ is given by

$$K(t,r) = \int_0^1 \cos\left[\left(3z^2-1\right)\omega_\perp(r)t\right]\mathrm{d}z \tag{9.3.6}$$

Here, we have substituted $\cos\theta$ by $z, \sin\theta\mathrm{d}\theta$ by $-\mathrm{d}\cos\theta$ and reversed direction of the integration, which compensated for the negative sign in $-\mathrm{d}\cos\theta$.

In practice, $D(t)$ is digitized and given as a vector at sampling times t_i. Likewise, it is sufficient to compute $P(r)$ as a vector at sampling distances r_k. The integral equation is thus transformed to a matrix equation

$$\vec{D} = \mathbf{K}\vec{P} \tag{9.3.7}$$

Unfortunately, this matrix equation cannot easily be inverted, since the rows of kernel $\mathbf{K}$ are not orthogonal, i.e., the scalar product of dipolar evolution function vectors at different r_k is not zero. The weak linear dependence of the rows makes the problem ill-posed. Small deviations of the experimental $\vec{D}$ from the "true" $\vec{D}_{\text{ideal}}$, for instance due to noise, cause large deviations of $\vec{P}$ from the true distance distribution. This problem can be solved only by taking into account additional information.

First, we know that, as a probability density, $P(r) \geq 0$ at all r. Hence, we can impose a non-negativity constraint on $\vec{P}$. It turns out that this is not sufficient for stabilizing the solution. Noise can be fitted by ragged distance distributions with many narrow peaks, although we know that the distance distribution must be smooth, as it arises from a continuous distribution of molecular conformations. Tikhonov regularization imposes a smoothness restraint by minimizing

$$G_\alpha = \rho + \alpha\eta \tag{9.3.8}$$

where

$$\rho = \|\mathbf{K}\vec{P} - \vec{D}\|\|^2 \tag{9.3.9}$$

is the mean square deviation between experimental and simulated data and

$$\eta = \left\|\hat{L}^{(2)}\vec{P}\right\|^2 \tag{9.3.10}$$

is the square norm of the second derivative, which can be computed from $\vec{P}$ by multiplication with the second derivative operator $\hat{L}^{(2)}$. The regularization parameter α determines the relative <figure>1 < $figcaption$ >2</figcaption></figure>

Copy and Paste
Image here.
Delete this
placeholder image

Figure 9.4: Tikhonov regularization of the data shown in Fig. 9.3. (a) L curve. The optimum regularization parameter corresponds to the corner (green circle) and provides the simulation shown in Fig. 9.3(b) as well as the extracted distance distribution shown as a black line in panel (c) of the current Figure. The red circle marks a too large regularization parameter that leads to oversmoothing. (b) Input form factor (black) and simulation for the too large regularization parameter corresponding to the red circle in the L curve. (c) Theoretical distance distribution used for simulating a noiseless form factor (green) and distance distribution extracted from the noisy form factor with optimum regularization parameter corresponding to the green circle in the L curve (black). (d) (c) Theoretical distance distribution used for simulating a noiseless form factor (green) and distance distribution extracted from the noisy form factor with a too large regularization parameter corresponding to the red circle in the L curve (black).

weight of the smoothing restraint with respect to mean square deviation between experimental and simulated data. A parametric plot of $\log \eta$ versus $\log \rho$ as a function of α is approximately L-shaped (Fig. 9.4). For very small α, roughness η of the distance distribution can be decreased strongly without increasing mean square deviation ρ very much. For large α, $\vec{P}$ is already smooth and a further increase of α will lead only to a small decrease in roughness η, but to a large increase in ρ, since the overly broadened distance distribution no longer fits the dipolar oscillations. Hence, in a mathematical sense, the optimum regularization parameter corresponds to the corner of the L curve. At this regularization parameter the extracted distance distribution (black line in Fig. 9.4(c)) is only slightly broader than the true distance distribution (green line) and the simulated form factor (red line in Fig. 9.3(b)) agrees with the experimental form factor (black line), except for the white noise contribution. If the regularization parameter is too large (red circle in Fig. 9.4(a)), the simulated form factor is overdamped (red line in Fig. 9.4(b)) and the distance distribution unrealistically broad (black line in Fig. 9.4(d)). For a too small regularization parameter the distance distribution unrealistically splits into several narrow peaks and the simulated form factor fits part of the noise (not shown). This error cannot be as clearly discerned in the simulated form factor as overdamping can be discerned. Undersmoothing is apparent only in the L curve.