Optimising NMR Spectroscopy through Method and Software Development

Jonathan Yong

University of Oxford

Contents

refsection:1

refsection:2

Chapter 2

Pure shift NMR

Pure shift NMR refers to the technique of acquiring NMR spectra free of multiplet structure, such that every chemical environment gives rise to a singlet.[1,2] In the context of this thesis, we use the term 'pure shift' exclusively to refer to broadband homodecoupled ^{1}H spectra. Here, 'broadband' means that the couplings are removed from the entire spectrum, as opposed to just a subset of it (which can be accomplished through band-selective refocusing). 'Homodecoupled' refers to the fact that the primary target here is the removal of homonuclear couplings, which cannot be done simply through decoupling during acquisition. Finally, although pure shift techniques can be applied to any nuclide, ^{1}H spectra are of greatest interest because of the narrow chemical shift range of ^{1}H which often leads to peak overlap. (Another prominent case is ^{13}C pure shift NMR in isotopically enriched molecules, but this is not within the scope of this thesis.)

In this chapter, I first cover the theory underpinning, and a brief history of, pure shift experiments. I then describe a variety of approaches aimed at increasing the quality of pure shift experiments: this is measured both in terms of *sensitivity* as well as *purity*, i.e. the lack of spectral artefacts arising from imperfect decoupling. In all cases, we compare these against the state-of-the-art PSYCHE pure shift method. Finally, I end with a section discussing the combination of pure shift diffusion spectroscopy—formally a pseudo-3D experiment—with the use of ultrafast (single-scan) NMR techniques to collapse the diffusion dimension. This project was carried out in collaboration with Jean-Nicolas Dumez (University of Nantes).

The work in this chapter has not been submitted for publication.

2.1 Theoretical background

In this section, I use a system with two weakly coupled spins I_1 and I_2 to illustrate the ideas behind pure shift NMR. This means that $H_{\text{free},I} = \Omega_1 I_{1z} + \Omega_2 I_{2z} + 2\pi J I_{1z} I_{2z}$; this is diagonal in

the Zeeman basis:

$$H_{\text{free}} = \begin{pmatrix} \omega_{\alpha\alpha} & 0 & 0 & 0 \\ 0 & \omega_{\alpha\beta} & 0 & 0 \\ 0 & 0 & \omega_{\beta\alpha} & 0 \\ 0 & 0 & 0 & \omega_{\beta\beta} \end{pmatrix}, \tag{2.1}$$ _{eq:h_free_weak}

where $\omega_{\lambda\mu} = \langle \lambda\mu | H_{\text{free},I} | \lambda\mu \rangle$ ($\lambda, \mu \in \{\alpha, \beta\}$) represents the precession frequency of the state $|\lambda\mu\rangle$. Given that $I_z |\alpha\rangle = (1/2) |\alpha\rangle$ and $I_z |\beta\rangle = -(1/2) |\beta\rangle$, these are relatively easy to work out:

$$\omega_{\alpha\alpha} = \frac{1}{2}(\Omega_1 + \Omega_2 + \pi J) \quad \omega_{\alpha\beta} = \frac{1}{2}(\Omega_1 - \Omega_2 - \pi J)$$
$$\omega_{\beta\alpha} = \frac{1}{2}(-\Omega_1 + \Omega_2 - \pi J) \quad \omega_{\beta\beta} = \frac{1}{2}(-\Omega_1 - \Omega_2 + \pi J). \tag{2.2}$$ _{eq:state_precessions}

The corresponding propagator for a time τ, $U = \exp(-iH_{\text{free},I}\tau)$, is then just:

$$U = \begin{pmatrix} \exp(-i\omega_{\alpha\alpha}\tau) & 0 & 0 & 0 \\ 0 & \exp(-i\omega_{\alpha\beta}\tau) & 0 & 0 \\ 0 & 0 & \exp(-i\omega_{\beta\alpha}\tau) & 0 \\ 0 & 0 & 0 & \exp(-i\omega_{\beta\beta}\tau) \end{pmatrix}. \tag{2.3}$$ _{eq:u_free_weak}

In the previous chapter, we showed how pulse sequences could be analysed using two different bases for spin-½ systems, depending on which was most mathematically expedient. To analyse pure shift NMR it turns out to be most convenient to introduce a third basis, namely $\{I_\alpha, I_\beta, I_+, I_-\}$.[3–5] The definitions of these were given in eq. (1.6): it is clear from there that these represent single matrix elements of ρ when expressed in the Zeeman basis. The same is true of their products in systems containing multiple spins. Consequently, the evolution of these operators under $H_{\text{free},I}$ is extraordinarily simple to calculate in matrix form: for example, we have that

$$I_{1+}I_{2\alpha} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.4}$$ _{eq:single_elem_plusalpha}

so $U(\tau)I_{1+}I_{2\alpha}U^\dagger(\tau)$ is simply

$$\begin{pmatrix} 0 & 0 & \exp(-i\omega_{\alpha\alpha}\tau)\exp(i\omega_{\beta\alpha}\tau) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \exp[-i(\Omega_1 + \pi J)\tau]I_{1+}I_{2\alpha}. \tag{2.5}$$ _{eq:single_elem_plusalpha}

Essentially, under $H_{\text{free},I}$, all of these operators merely acquire a phase factor which depends on the difference between two of the frequencies. The rules for the single-quantum operators on

spin 1 are explicitly given here:[*]

$$I_{1+}I_{2\alpha} \longrightarrow \exp[-i(\Omega_1 + \pi J)\tau]I_{1+}I_{2\alpha} \tag{2.6}$$ {eq:shift_basis_evolut

$$I_{1+}I_{2\beta} \longrightarrow \exp[-i(\Omega_1 - \pi J)\tau]I_{1+}I_{2\beta} \tag{2.7}$$ {eq:shift_basis_evolut

$$I_{1-}I_{2\alpha} \longrightarrow \exp[i(\Omega_1 + \pi J)\tau]I_{1-}I_{2\alpha} \tag{2.8}$$ {eq:shift_basis_evolut

$$I_{1-}I_{2\beta} \longrightarrow \exp[i(\Omega_1 - \pi J)\tau]I_{1-}I_{2\beta} \tag{2.9}$$ {eq:shift_basis_evolut

The rules for the corresponding operators on spin 2 can be easily obtained by permutation of labels. Notice that the evolution frequencies of the -1-quantum operators each correspond to one peak of the multiplet: for example, $\Omega_1 + \pi J$ and $\Omega_1 - \pi J$ (eqs. (2.8) and (2.9)) correspond to the two peaks of the spin-1 doublet.

Consider now a simple spin echo sequence: $90°_x$–τ–$180°_x$–τ–detection. The initial excitation pulse acts on both spins 1 and 2, and thus generates a mixture of all eight possible single-quantum operators (the four above plus four more on spin 2). For simplicity, we only consider the $I_{1+}I_{2\alpha}$ term. This evolves in the first τ delay to give $\exp[-i(\Omega_1 + \pi J)\tau]I_{1+}I_{2\alpha}$. The 180° pulse *flips* both spins 1 and 2, in that it causes the transitions $I_+ \leftrightarrow I_-$ and $I_\alpha \leftrightarrow I_\beta$; consequently, we have that

$$\exp[-i(\Omega_1 + \pi J)\tau]I_{1+}I_{2\alpha} \longrightarrow \exp[-i(\Omega_1 + \pi J)\tau]I_{1-}I_{2\beta}. \tag{2.10}$$ {eq:spin_echo_1}

In the second delay, we get a second phase factor from the evolution of the $I_{1-}I_{2\beta}$ operator:

$$\exp[-i(\Omega_1 + \pi J)\tau]I_{1-}I_{2\beta} \longrightarrow \exp[-i(\Omega_1 + \pi J)\tau]\exp[i(\Omega_1 - \pi J)\tau]I_{1-}I_{2\beta}$$
$$= \exp(-2i\pi J\tau)I_{1-}I_{2\beta}. \tag{2.11}$$ {eq:spin_echo_2}

Detection of this gives us one of the two peaks of the spin-1 doublet, as described previously, but with a phase factor tacked on. The Ω_1 terms in the phase factor are cancelled out, which reflects the fact that the offset (or chemical shift) is refocused by the 180° pulse. However, the J-evolution is not refocused, which leads to characteristic phase distortions in the detected multiplets. The same is true of all the other single-quantum operators on both spins.

In order to refocus the J-evolution as well as the chemical shift, we would need—instead of a 180° pulse—a pulse sequence element which simultaneously effects *all* of the following transitions:

$$I_{1+}I_{2\alpha} \longrightarrow I_{1-}I_{2\alpha}; \quad I_{1+}I_{2\beta} \longrightarrow I_{1-}I_{2\beta}; \quad I_{1-}I_{2\alpha} \longrightarrow I_{1+}I_{2\alpha}; \quad I_{1-}I_{2\beta} \longrightarrow I_{1+}I_{2\beta}; \tag{2.12}$$ {eq:pure_shift_require

$$I_{1\alpha}I_{2+} \longrightarrow I_{1\alpha}I_{2-}; \quad I_{1\beta}I_{2+} \longrightarrow I_{1\beta}I_{2+}; \quad I_{1\alpha}I_{2-} \longrightarrow I_{1\alpha}I_{2+}; \quad I_{1\beta}I_{2-} \longrightarrow I_{1\beta}I_{2-}. \tag{2.13}$$ {eq:pure_shift_require

---

[*]Note that § 10.4.2 of Keeler's text[3] has a sign error in these equations—the sign of the πJ evolution is flipped—though it proves to be inconsequential as the sign of J cannot be observed.

Such an element forms the basis of a pure shift technique, and I refer to it here as a *pure shift element* (PSE). The difficulty in designing a PSE is that *all* spins must be simultaneously decoupled from each other (and not just one spin). For example, if we only had to invert spin 1 and not spin 2 (i.e. only eq. (2.12) and not eq. (2.13)), this could be trivially accomplished by a selective 180° pulse on spin 1. However, this would not bring about the correct transitions for spin 2. Yet another complicating factor is that the spins will have offsets and couplings which are *a priori* not known; so the design of the PSE cannot use such parameters as inputs. These limitations mean that it is impossible to accomplish the above transitions *in full*; rather, a more realistic scenario involves

$$I_{1+}I_{2\alpha} \longrightarrow cI_{1-}I_{2\alpha} + \sum_i c_i' M_i \qquad (2.14) \quad \text{\{eq:realistic\_pse\}}$$

and likewise for the other operators. Here, the desired transition probability c directly correlates with the sensitivity of the PSE, and the $M_i's$ are some other undesired operators which (if detectable) lead to artefacts if not suppressed.

In the above discussion, note that the role of the PSE is to invert the I_+ and I_- terms, and to retain the I_α and I_β terms. The prevailing terminology is to refer to the spins with I_+ and I_- terms as *active spins*, and the I_α and I_β spins as *passive spins*. Thus, for example, in the context of eq. (2.12), spin 1 is active and spin 2 is passive. The detected signal always arises from the active spins.

Before moving on to the discussion of how this is accomplished in practice, I insert a slight digression about *J-resolved* (or *2DJ*) *spectroscopy*, which is very closely related to pure shift NMR. The basic 2DJ sequence involves a spin echo of duration t_1 (i.e. the τ delay above is $t_1/2$), immediately followed by detection. If we only consider a single operator and reuse the analysis above, then we have that

$$\begin{aligned}
I_{1+}I_{2\alpha} &\xrightarrow{t_1/2} \exp[-\mathrm{i}(\Omega_1 + \pi J)t_1/2]I_{1+}I_{2\alpha} \\
&\xrightarrow{180°} \exp[-\mathrm{i}(\Omega_1 + \pi J)t_1/2]I_{1-}I_{2\beta} \\
&\xrightarrow{t_1/2} \exp(-\mathrm{i}\pi J t_1)I_{1-}I_{2\beta}.
\end{aligned} \qquad (2.15) \quad \text{\{eq:2dj\_operator\_analy\}}$$

This yields a complex signal of the form

$$s(t_1, t_2) = \exp(-\mathrm{i}\pi J t_1)\exp[\mathrm{i}(\Omega_1 - \pi J)t_2], \qquad (2.16) \quad \text{\{eq:2dj\_signal\}}$$

which when Fourier transformed yields a phase twist lineshape at $(-\pi J, \Omega_1 - \pi J)$. The other component on spin 1 (starting from $I_{1+}I_{2\beta}$) likewise yields a phase twist at $(\pi J, \Omega_1 + \pi J)$. It has long been known that *shearing* this 2DJ spectrum by 45° (i.e. moving each data point (Ω_1, Ω_2) to $(\Omega_1, \Omega_2 - \Omega_1)$) generates a spectrum which only has chemical shift information in the ω_2 dimension. After magnitude-mode processing, projection of this spectrum onto the ω_2 axis, for

example, would in principle yield a pure shift spectrum.[*] This is true, but in practice the phase twist lineshapes cause the resulting resolution to be very poor, which defeats the purpose of using a pure shift spectrum. To circumvent this issue, a glut of special processing techniques have been proposed[7−9] (see also references therein); but more ideally, we want a phase-sensitive 2DJ spectrum, where a pair of 'echo' and 'antiecho' signals are obtained:[†]

$$s_{\text{echo}}(t_1, t_2) = \exp(-i\pi J t_1) \exp[i(\Omega_1 - \pi J)t_2] \tag{2.17}$$ `{eq:2dj_signal_ea}`

$$s_{\text{antiecho}}(t_1, t_2) = \exp(i\pi J t_1) \exp[i(\Omega_1 - \pi J)t_2] \tag{2.18}$$

and processed in the same way as previously described in § 1.4.3 to yield the double absorption-mode lineshape. The echo signal is of course the same as in eq. (2.16), but to obtain the antiecho signal, we require a different pulse sequence with a PSE inserted right before detection. Note that we start with a different operator here, $I_{1-}I_{2\alpha}$, in order to end up with the same $I_{1-}I_{2\beta}$ operator just before detection.

$$I_{1-}I_{2\alpha} \xrightarrow{t_1/2} \exp[i(\Omega_1 + \pi J)t_1/2]I_{1-}I_{2\alpha} \tag{2.19}$$ `{eq:2dj_signal_antiech}`

$$\xrightarrow{180°} \exp[i(\Omega_1 + \pi J)t_1/2]I_{1+}I_{2\beta} \tag{2.20}$$

$$\xrightarrow{t_1/2} \exp(i\pi J t_1)I_{1+}I_{2\beta} \tag{2.21}$$

$$\xrightarrow{\text{PSE}} c\exp(i\pi J t_1)I_{1-}I_{2\beta}. \tag{2.22}$$

In order to apply echo–antiecho processing, the decrease in sensitivity by a factor of c must also be applied to the echo spectrum: this can be done by simply inserting the PSE before the t_1 spin echo, which does not affect the relative modulation in t_1 and t_2. Thus, we see that *exactly the same PSE* allows us to generate pure shift spectra as well as absorption-mode 2DJ spectra, a fact which has been previously demonstrated using various PSEs.[10,11] In fact, the same formalism can be used to describe a family of small flip angle COSY experiments, including ECOSY[5,12,13] and *z*-COSY;[14−16] these are also closely related to pure shift NMR. In particular, the anti *z*-COSY experiment is a precursor to PSYCHE, and is analysed in § 2.2.3. However, a full discussion of these is beyond the scope of this thesis.

---

[*]Or equivalently, projection of the unsheared spectrum along a 45° axis.[6]

[†]The terms 'echo' and 'antiecho' refer to the relative senses of the coherences evolving during t_1 and t_2: in the echo spectrum these have opposite signs, e.g. $I_z S_+$ and I_- in the HSQC, and in the antiecho spectrum they have the same sign. As pointed out by Pell and Keeler,[10] this is not really appropriate for the 2DJ experiment since each half of t_1 has a coherence with a different sense, but we will stick to this nomenclature as the underlying concept is very similar to that of echo–antiecho processing.

2.2 Pure shift in practice

The previous section described the underlying theory used for analysing PSEs and showed how such an element could be used to record absorption-mode 2DJ spectra. From this, one can obtain a pure shift spectrum through shearing and projection. However, this is only an *indirect* route to a pure shift spectrum. In this section, we will tackle the main question of how pure shift experiments may be *directly* acquired. Following this, I cover several examples of PSEs reported in the literature. This selection is not exhaustive: I only choose to cover a handful of PSEs which *specifically* accomplish the transformations listed in eqs. (2.12) and (2.13).

2.2.1 Acquisition modes
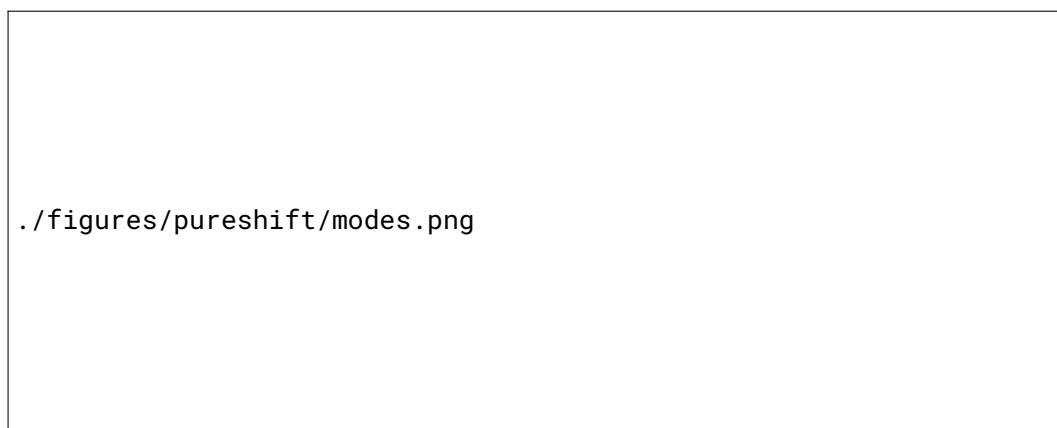
Figure 2.1: Possible acquisition modes for pure shift spectroscopy. The red box labelled 'PSE' indicates a generic pure shift element, which can be any of those described in the main text. In practice, gradients are also used to suppress unwanted coherence transfers; these are not shown here for simplicity. **(a)** Insertion of a J-refocusing element (JRE) in the centre of an indirect-dimension evolution period, which leads to a spectrum which is pure shift in F_1. The 90°–τ_m–90° mixing period shown here is that of a NOESY experiment, but in principle it can be anything. **(b)** Real-time acquisition of a 1D pure shift spectrum in chunks of duration T_{chunk}. **(c)** Interferogram acquisition of a 1D pure shift spectrum, where t_1 is lengthened by T_{chunk} every increment.

Restating eq. (2.14), suppose we have a PSE which accomplishes the transformation

$$I_{1+}I_{2\alpha} \longrightarrow cI_{1-}I_{2\alpha} + \sum_i c_i' M_i. \qquad (2.23)$$

The simplest method of using this is to insert it in the middle of a t_1 period of a 2D experiment. This is actually not entirely desirable, because the PSE causes *both* chemical shifts and J-couplings to be refocused; consequently, there will be *no* frequency modulation during t_1 at all! It is more sensible to combine the PSE with a hard 180° pulse (which refocuses only chemical shifts). Together, the effect is to refocus J-couplings and allow chemical shifts to evolve; this combination

is thus called a *J-refocusing element*, or JRE (fig. 2.1a). We can equivalently say that the JRE flips all passive spins and leaves active spins untouched. This distinction between a JRE and a PSE is important and will be referred to several times in this chapter.

Implementing a JRE in the middle of a t_1 period is simple, requiring minimal modification of existing 2D experiments. Furthermore, the pure shift 'character' of the F_1 dimension may then be mapped to the F_2 dimension through indirect covariance processing.[17–22] However, the increased resolution in the F_1 dimension provided by homodecoupling cannot really be reaped unless many t_1 increments are acquired. Furthermore, this does not help with acquiring a 1D pure shift spectrum, where there is no indirect dimension.

During direct detection, if coherences are allowed to evolve at their 'intrinsic' frequencies, no decoupling can be accomplished. In order to effectively utilise a JRE in a 1D experiment, or the direct dimension of a 2D experiment, it is necessary to periodically interrupt the acquisition and insert a JRE: this causes the chemical shift evolution to be effectively 'suspended' for the duration of the JRE, and the sense of J-evolution to be reversed (fig. 2.1b).[⋆] This leads to a series of FID 'chunks' which must then be concatenated to form the desired FID; the required spacing of the JREs, or equivalently the duration of each chunk T_{chunk}, must satisfy $T_{\text{chunk}} \ll 1/J$ (in practice, it is on the order of $1/(2J)$). This is referred to as real-time acquisition,[23–25] and although one still has to pay the sensitivity price of c, it allows a pure shift spectrum to be acquired in effectively the same time as the original coupled spectrum. Its 'single-scan' nature also allows, for example, the application of hyperpolarisation techniques which cannot be reproducibly repeated.[26,27]

Unfortunately, it is not always possible to perform real-time acquisition. The reason is because the JRE is applied multiple times, and each time it is, it must select for the same active and passive spins in the same molecule as it did the last time. In other words, *for any given molecule in the sample*, it must enforce this CTP:

$$I_{1-}I_{2\alpha} \xrightarrow{\text{JRE}} I_{1-}I_{2\beta} \xrightarrow{\text{JRE}} I_{1-}I_{2\alpha} \xrightarrow{\text{JRE}} I_{1-}I_{2\beta} \xrightarrow{\text{JRE}} \cdots \qquad (2.24)$$

{eq:real_time_pureshif

As will be described later, the BIRD and Zangger–Sterk methods always select the same active spins in the same molecules, but the PSYCHE method notably does not. Therefore, in order to acquire pure shift PSYCHE spectra, we have to resort to the *interferogram method*, where each chunk is obtained as a separate increment of a 2D experiment (fig. 2.1c). The insertion of the JRE in the middle of the t_1 period means that when detection is started, it is 'as if' only the chemical shift has evolved for a period t_1. On each increment, one chunk—again of duration

---

[⋆]Relaxation during the JRE must also be taken into account for the real-time method: this causes each successive chunk to decay in intensity faster than usual, thereby leading to peak broadening, which can be an issue for very long JREs. In contrast, the interferogram method as only one JRE is applied on each increment, so the losses due to relaxation during the JRE are simply a constant factor.

$T_{\text{chunk}} \ll 1/J$—is detected, and then t_1 is incremented by T_{chunk} so that the next chunk can be recorded. Finally, the chunks are stitched together to form the requisite FID.[*] Since the indirect dimension is not processed by a Fourier transform, this is sometimes called a *pseudo-2D* experiment (or *pseudo-3D* if this is applied to the direct dimension of a 2D experiment, and so on). In this case, the sensitivity drop c is incurred, but there is on top of that also a time penalty in that the experiment duration must be lengthened by n times in order to collect n chunks. n is typically on the order of 16–32.

2.2.2   Pure shift elements

Figure 2.2: A selection of pure shift elements. **(a)** Zangger–Sterk PSE,[30] involving the combination of a selective 180° refocusing pulse and a weak gradient. **(b)** BIRD PSE;[31,32] the delay Δ is set to $1/(4 \cdot {}^1J_{\text{CH}})$. **(c)** Time-reversal PSE,[13] simply consisting of a hard pulse with variable flip angle β. Multiple spectra with different values of β must be co-added to suppress artefacts (though not completely, as discussed in the text). **(d)** PSYCHE PSE,[33] consisting of two saltire pulses[34,35] with flip angle β, and a weak gradient.

We are finally now in a position to describe individual PSEs and their mechanism of action. We begin with the Zangger–Sterk (ZS, or 'slice-selective') PSE,[30] where a selective refocusing pulse and a weak gradient are simultaneously applied (fig. 2.2a). In practice, an rSNOB pulse[36] is often used as the refocusing pulse. The effect of the gradient is to make each spin in the sample have a spatially dependent offset; conversely, since the the refocusing pulse has a specific, narrow, bandwidth, in each *slice* (or cross-section) of the sample a different spin will fall within this bandwidth. This spin is refocused by the PSE and therefore becomes the active spin *within that specific slice*; the bracketing pair of CTP gradients serve to destroy coherences on all the other spins which are not inverted. Each signal of the pure shift spectrum therefore derives from a specific slice of the sample; during direct detection, all slices simultaneously contribute to the signal, thus yielding a broadband pure shift spectrum.

The sensitivity of the ZS method tends to be low (the factor c tends to be on the order of ~ 0.01 to 0.05), as each signal only derives from a narrow region of the sample. Nevertheless, it still finds wide usage in pure shift applications nowadays, especially because it is compatible with

---

[*]Not included in fig. 2.1c is the extra detail that scalar couplings are usually allowed to evolve for a period of $T_{\text{chunk}}/2$ at the start of the sequence, by the addition of a spin echo: this amounts to a *prefocusing* of J-evolution, such that J-coupling is refocused in the middle of the chunk rather than the beginning.[28] This allows chunk sizes twice as large to be used, reducing the total duration of the experiment. This J-prefocusing can also be done in a more intelligent manner via the SAPPHIRE method,[29] which is discussed in more detail in § 4.4.4.

the real-time acquisition mode:[24] as long as the pulse and the weak gradient are the same each time, then the same active spins will always be chosen in the same slice (as long as diffusion effects are ignored). The PSE can also be customised through the bandwidth of the refocusing pulse: decreasing this improves the spatial differentiation between spins which have similar intrinsic offsets, yielding better decoupling quality, albeit at the cost of sensitivity. The ZS element can be easily—and has been—adapted for use in many experiments, including (but not limited to) absorption-mode 2DJ spectroscopy[10] and selective refocusing (SERF) experiments for the measurement of $^{n}J_{\mathrm{HH}}$.[37–40]

Next up is the *bilinear rotation decoupling* (BIRD) pulse element (fig. 2.2b). BIRD is not spatially selective like the ZS method; instead, it is *isotope-selective* in that it acts as a 180°_{y} pulse on ^{13}C–bound protons, and does not affect ^{12}C–bound protons. Consequently, all ^{13}C–bound protons become the active spins in the context of pure shift NMR. The first report of the BIRD element,[31] in 1982, was clearly ahead of its time: it reported the use of an interferogram-type approach to obtain 1D pure shift spectra. However, in the subsequent decades, this seemed to have been forgotten: BIRD found much more use as an isotope-selective rotation element in heteronuclear NMR,[41] until its use as a pure shift element was 'rediscovered'.[32,42]

An immediate drawback of BIRD is that it does not decouple geminal (diastereotopic) CH_2 groups, as both protons would be either both active or both passive. The sensitivity penalty of BIRD is also relatively severe: the factor c derives from the natural abundance of ^{13}C, which is approximately 0.011. However, it is also compatible with real-time acquisition,[23] and has found particular success as a pure shift element in the F_2 dimension of HMQC and HSQC experiments:[25,42–46] in this case, the use of BIRD leads to no loss of sensitivity as only ^{13}C–bound protons are detected in HSQC experiments anyway.* It should be noted that the BIRD element does not need to be combined with a hard $180°$ pulse to form a JRE: the inverse effect of only flipping the passive spins can be accomplished by simply changing the phase of both internal $180°$ pulses to $+y$. Using the nomenclature of Uhrín et al.,[41] the PSE and JRE forms of the BIRD pulse can be labelled as $\mathrm{BIRD}^{\mathrm{d,X}}$ and $\mathrm{BIRD}^{\mathrm{r,X}}$ respectively.

The *time-reversal* PSE (fig. 2.2c) is even simpler: it consists only of one hard pulse with flip angle β.[13] The catch is that the experiment must be repeated with different values of β, and the results added up with specific weightings.[5] This leads to cancellation of *some*, but not all, of the unwanted coherences: for example, on its own, it does not suppress COSY-type coherence transfer such as $I_{1-}I_{2\alpha} \rightarrow I_{1\alpha}I_{2+}$. This proved to be inconsequential in the original application of F_1 decoupling in a 2D NOESY;[13] however, it is not acceptable in a 1D context as it leads to substantial artefacts. This will be discussed in more detail in § 2.5, so a fuller analysis of the time-reversal PSE is deferred until then.

---

*In fact, the sensitivity is increased by the collapse of multiplet structure.

Finally, we come to the PSYCHE method, which is generally considered the current state of the art for pure shift NMR.[33] The corresponding PSE consists of two swept-frequency pulses applied during a weak gradient: for example, a pair of chirps with opposite sweep directions can be used. Using two *saltire pulses*, pulses which simultaneously sweep in both directions, leads to an increased signal-to-artefact ratio.[34] The operation of this PSE is not easy to fully explain.[35] However, we can adopt the (not fully accurate, but still useful) instant-flip approximation[47,48]—that the swept-frequency pulse acts as an instantaneous 180° rotation on each frequency it passes through. Using this, the PSYCHE element (or strictly, the PSYCHE JRE, including a hard 180° pulse) may be viewed as a spatially parallelised version of the anti z-COSY experiment,[14,15] which we now describe.

2.2.3 PSYCHE in detail

Figure 2.3: A closer look at the mechanism of the PSYCHE PSE. **(a)** The β–z-filter–β mixing period used in the original anti z-COSY experiment. This has a similar action to a JRE, but does not suppress COSY-type coherence transfers from spin i to spin j. **(b)** An illustration of how COSY-type artefacts are suppressed by the PSYCHE pulse element. The desired CTP which remains on spin 1 is rephased by the gradients, but the COSY CTP is dephased. **(c)** An illustration of how zero-quantum terms are suppressed by the PSYCHE element through spatial averaging: in each slice of the sample (highlighted in blue and orange), the zero-quantum terms are allowed to evolve for a different duration.

The anti z-COSY experiment utilises a β–z-filter–β mixing period (fig. 2.3a), where β is a small angle, typically 10° to 20°. The role of the z-filter[49] is to remove zero-quantum terms such as $I_{1-}I_{2+}$ between the two β pulses, retaining only population terms such as $I_{1\alpha}I_{2\alpha}$. In an elegant paper by Pell et al.,[15] it was shown that by isolating the diagonal peaks of a 2D anti z-COSY experiment and taking the 45° projection of these, a pure shift spectrum could be obtained. Here, we will go one step further and consider the *direct* use of this element as a JRE: this will reveal problems which are circumvented by the PSYCHE experiment.

We first need to introduce how the basis operators $\{I_+, I_-, I_\alpha, I_\beta\}$ evolve under a hard pulse (applied

along the *x*-axis) with flip angle β:

$$I_{\pm} \xrightarrow{\beta I_x} c^2 I_{\pm} + s^2 I_{\mp} \pm \frac{iS}{2}(I_{\alpha} - I_{\beta}) \tag{2.25}$$

{eq:ipm_beta_evolution

$$I_{\alpha} \xrightarrow{\beta I_x} c^2 I_{\alpha} + s^2 I_{\beta} + \frac{iS}{2}(I_+ - I_-) \tag{2.26}$$

{eq:ialpha_beta_evolut

$$I_{\beta} \xrightarrow{\beta I_x} c^2 I_{\beta} + s^2 I_{\alpha} - \frac{iS}{2}(I_+ - I_-) \tag{2.27}$$

{eq:ibeta_beta_evoluti

where $S = \sin\beta$, $s = \sin(\beta/2)$, and $c = \cos(\beta/2)$. Using these formulae, an analysis on a two-spin system (see Pell et al. for the equivalent analysis on a three-spin system) shows that this element converts the term $I_{1+}I_{2\alpha}$ to

$$I_{1+}I_{2\alpha} \longrightarrow \underbrace{\frac{1}{2}S^2c^4 I_{1+}I_{2\beta}}_{\text{term 1}} + \underbrace{S^2c^2s^2 I_{1+}I_{2\alpha}}_{\text{term 2}} - \underbrace{\frac{1}{4}S^2c^4 I_{1\alpha}I_{2+} + \frac{1}{4}S^2c^4 I_{1\beta}I_{2+}}_{\text{terms 3 and 4}} + \cdots, \tag{2.28}$$

{eq:anti_z_cosy_transi

where other terms with different coherence orders have been neglected (on the basis that they can be easily suppressed with bracketing CTP gradients), and terms with higher orders in *s* have been discarded since $s = \sin(\beta/2) \ll 1$ for small β.

The first term $I_{1+}I_{2\beta}$, corresponding to the flipping of passive spins only, is the only term we want to see from a JRE. The second term, $I_{1+}I_{2\alpha}$, corresponds to the case where neither active nor passive spins have been flipped. In the original anti *z*-COSY work, these give rise to 'off-diagonal' peaks which are part of a multiplet on the diagonal, but when projected at 45° generate artefacts around the pure shift peak. In the context of pure shift NMR, these are called 'recoupling artefacts', as they arise from imperfect J-refocusing. Note that the ratio of recoupling artefacts to desired signal is proportional to $S^2c^2s^2/S^2c^4 = \tan^2(\beta/2)$: using a smaller value for β therefore leads to better signal-to-artefact ratios, but also lower overall sensitivity. The PSYCHE element is similar to the β–*z*-filter–β element in this regard: it does not suppress the recoupling artefacts, but instead relies on the user choosing a suitable value for β such that the artefact-to-signal ratio is tolerably small. If the sensitivity proves to be insufficient, the flip angle β may be increased instead: this leads to a larger artefact-to-signal ratio, but if the sample is not concentrated anyway, it may well be that the artefacts do not rise above the noise level.

The third and fourth terms, $I_{1\lambda}I_{2+}$ ($\lambda \in \{\alpha, \beta\}$), represent 'COSY-type' coherence transfer to a coupled spin: these led to crosspeak multiplets in the original anti *z*-COSY, which could be removed by hand before taking the projection. However, in a pure shift sequence, the peaks arising from these terms cannot simply be removed in the same way. It is precisely this issue which precludes the β–*z*-filter–β element from being directly used as a JRE, and motivates the development of the PSYCHE PSE, which *does* suppress these coherence transfers.

To understand how this occurs, we invoke the instantaneous spin-flip assumption. Each coher-

ence I_{i+} is converted (or 'flipped') to a population term $I_{1\lambda}$ at a specific point $\alpha\tau_{\mathrm{p}}$ after the start of the first chirp, and can be reconverted to a coherence on the same spin I_{i-} at a time $\alpha\tau_{\mathrm{p}}$ before the end of the second chirp (the green CTP in fig. 2.3b).* Here, τ_{p} is the duration of the chirp and $0 < \alpha < 1$. In this case, the coherence is perfectly rephased by the weak gradient applied during the chirp pulses, since the *time* it experiences these gradients for is the same on both sides of the chirps. Now, if the I_{i+} term is converted to a coherence on a different spin I_{j-}, it experiences the gradient for a total duration of $\alpha\tau_{\mathrm{p}}$ after the start of the first chirp, and also $\alpha'\tau_{\mathrm{p}}$ before the end of the second chirp (the red CTP in fig. 2.3b). In general, $\alpha \neq \alpha'$ as spins i and j have different offsets; therefore, this CTP is dephased by the gradients, resulting in complete suppression of the COSY-type artefacts in the spectrum.

It remains to also consider how the PSYCHE element selects for the population terms between the two spin flips. Any terms with nonzero coherence order are simply dephased by the weak gradient. However, zero-quantum terms (in homonuclear systems) are not dephased by gradients, and to eliminate them in a single-scan manner, they must be spatially averaged: this is what the *z*-filter in the anti *z*-COSY does, for example. It turns out that the PSYCHE element also does a similar spatial averaging. Following on from the previous paragraph, the time *between* the spin flips (for the desired pathways, i.e. not COSY-type coherence transfer) is given by $2(1 - \alpha)\tau_{\mathrm{p}}$. At the same time, the weak gradient induces a range of offsets across the sample, much like in the Zangger–Sterk experiment. Thus, the offset, and thus the value of α, for a given spin depends on which slice it is in; for example, fig. 2.3c uses values of α_1 and α_2 for two different slices (blue and orange). If zero-quantum terms are present between the spin flips, they evolve during this time and accrue a spatially-dependent phase: summation of these during FID acquisition leads to a cancellation of these terms. Only population terms such as $I_{1\alpha}I_{2\alpha}$ survive during this, as they do not precess during this time.

The sensitivity of PSYCHE is significantly better than for other methods: depending on the flip angle β chosen, c is typically on the order of 0.05–0.15 (see also fig. 2.4 for explicit simulations). Furthermore, it is generally more robust with respect to strong coupling compared to other pure shift methods (artefacts from strong coupling often arise due to unexpected coherence transfer,[4] which is suppressed in a similar way to the COSY-type artefacts). These two factors alone have meant that PSYCHE has enjoyed substantial adoption since its introduction: a large number of 2D experiments utilising PSYCHE decoupling in either F_1 or F_2 have been developed,[22,50–55] notably the PSYCHE-iDOSY diffusion experiment,[56] where the increased resolution provided by pure shift spectroscopy translates into increased resolution in the *diffusion* dimension as well. Like the ZS element before it, the PSYCHE element has also been used for the acquisition of absorption-mode 2DJ spectra.[11]

---

*Note the change in the sign of the coherence, which differs from the analysis of the anti *z*-COSY experiment. This arises because we are only considering the PSYCHE PSE on its own, *not* the JRE.

Despite this success, PSYCHE suffers from one significant drawback: it cannot be used in a real-time fashion. The PSYCHE PSE is often said to select active and passive spins in a 'statistical' manner: this is because of the c^2 and s^2 terms arising from the low-flip angle pulses. What this really means is that we do not care *exactly* which spins are active and which are passive, but that a certain proportion of the spins are active and passive. Repeated application of the PSE therefore does not select for the same active spins each time, which precludes its application to real-time acquisition.

Although the PSYCHE PSE may appear deceptively simple at first glance, the closer analysis given here (and elsewhere[34]) clearly shows that its inner workings are anything but. Along with other ingenious experiments such as the *z*-filter,[4] ultrafast NMR,[57–59] and more recently GEMSTONE,[60] PSYCHE is a prime example of how *spatiotemporal averaging* and pulsed field gradients can be used to great effect in modern NMR spectroscopy.[61]

At the same time, PSYCHE itself is not *perfect*: it does not fully suppress recoupling artefacts, and can only be used in the interferogram mode. To improve on PSYCHE would therefore entail one of the following:

1. increasing the sensitivity (while maintaining purity);

2. increasing the purity (while maintaining sensitivity); or

3. developing a pure shift element which is compatible with real-time acquisition while giving comparable sensitivity and purity to PSYCHE.

The sections which follow describe my efforts towards objectives (1) and (2).

2.3 PSYCHE with a variable number of saltires

The first attempted method was to change the number of saltire pulses used in the waveform. As described above, PSYCHE relies on spatiotemporal averaging to suppress unwanted artefacts: this crucially relies on the fact that the pulse(s) used in the PSE are symmetric. This can be accomplished with two opposing chirps, or two saltires, both of which are symmetric. However, a *single* saltire is also symmetric in itself: it is not hard to show that the analysis given in § 2.2.3 is still valid for a single saltire. Likewise, the use of four saltires would also be valid.

Theoretical simulations show that the overall profile of signal and artefact versus flip angle (unsurprisingly) varies with the number of saltires: in particular, using more saltires with a smaller flip angle generally accomplishes a similar effect (fig. 2.4). This may be qualitatively rationalised as more saltires providing more possible CTPs which generate the desired signal (the same idea is generally invoked when discussing the difference between unidirectional chirps

fig:fa_dependence_124

Figure 2.4: Simulated signal and artefact intensity for various PSYCHE-style PSEs as a function of the flip angle used. Calculations were performed on a two-spin system with a coupling of 7 Hz, and an offset difference of 1.5 kHz. The total PSE duration was 30 ms (so, for example, in the four-saltire PSYCHE, each saltire was 7.5 ms). Solid lines indicate the coefficients for the desired $I_{1+}I_{2\alpha} \to I_{1-}I_{2\alpha}$ pathway which contributes to the pure shift signal; dashed lines the coefficients for the undesired $I_{1+}I_{2\alpha} \to I_{1-}I_{2\beta}$ pathway, which gives rise to recoupling artefacts.

and saltires[34]). However, regardless of the number of saltires, the fundamental strategy of adjusting the flip angle to control the signal-to-artefact ratio remains valid, which naturally raises the question of whether specific pulse parameters can be chosen in order to obtain the best decoupling quality.

We first consider the quadruple-saltire PSYCHE. In the first instance, we set the total duration of the PSE to 30 ms: this implies that each saltire is 7.5 ms long. For this experiment, the sensitivity was defined as the maximum height of the main peak in fig. 2.5a, and the artefact as the mean of the maximum heights of the two artefacts surrounding it. The plot in fig. 2.5b shows how these quantities vary as a function of the flip angle. Interestingly, for the artefacts, the profile observed is similar to that in the simulations: the double-saltire version performs better at low and very high flip angles, but in the middle there is a region where the quadruple-saltire version has lower artefact intensity. The signal intensities for both the double- and quadruple-saltire versions, however, seem to plateau off rather more quickly than the simulations suggest.

To highlight one particular data point, fig. 2.5b suggests that the quadruple-saltire experiment with $\beta \approx 55°$ has a similar artefact level to the double-saltire experiment with $\beta \approx 20°$, but with

Figure 2.5: **(a)** The peak in the pure shift spectrum of andrographolide used for calculation of signal and artefact intensity. The recoupling artefacts flanking the main peak are clearly visible. **(b)** Peak heights of the desired signal (the central peak) and artefacts (the mean of the two flanking peaks), as a function of flip angle. **(c)** Quadruple-saltire PSYCHE with $\beta = 55°$. **(d)** The reference spectrum, a double-saltire PSYCHE with $\beta = 20°$. The peak at 3.85 ppm used for the sensitivity and purity analysis is labelled with an arrow. *Data code:* `7A-201016`.

substantially greater signal intensity. Insets from these two spectra are respectively shown in figs. 2.5c and 2.5d. In fact, this conclusion seems to be true for the peak in question, which is at the right edge of the spectral insets shown here. Overall, the quadruple-saltire 55° experiment does have better sensitivity than the double-saltire 20° experiment. However, the artefact intensity is not always suppressed as nicely as it is in the chosen reference peak: for example, the peak at 4.05 ppm is significantly less clean in the quadruple-saltire experiment. Any apparent advantage over the double-saltire experiment is therefore not very clear—at least from this data alone.[*]

---

[*]In fact, I also performed some preliminary experiments where the four-saltire PSE was lengthened to 60 ms. The artefact behaviour in these spectra were better than in their 30 ms counterparts, which is to be expected since a longer PSE leads to better spatiotemporal averaging. However, I unfortunately did not compare these against a 60 ms double-saltire experiment, so these results are not included in this thesis. (It would be rather unfair to compare them against the 30 ms double-saltire.)

Moving on to the single-saltire case, in addition to searching for a better signal and artefact profile, another possible motivation would be that the duration of the PSE can be decreased. This would minimise losses due to relaxation and diffusion during the PSE, which were not taken into account in fig. 2.4 (and the simulations there did not vary τ_p anyway). A single-saltire PSYCHE experiment was thus recorded with various combinations of flip angle ($\beta/° \in \{15, 20, 25, 28, 30\}$) and pulse duration ($\tau_p/\text{ms} \in \{15, 25, 30, 35, 45, 55\}$). In a similar way to the quadruple-saltire study, the sensitivity here was defined as the maximum height of the main peak in fig. 2.6a, and the artefact as the mean of the heights of the two artefacts surrounding it. (However, note that the compound used was different.) The purity, or signal-to-artefact ratio, is simply the former divided by the latter.[*]

Figure 2.6: **(a)** The peak in the (reference, double-saltire) pure shift spectrum of cyclosporin used for calculation of signal intensity and signal-to-artefact ratio. **(b)** Plot showing the signal intensity and signal-to-artefact ratio obtained in various single-saltire PSYCHE experiments, normalised against the double saltire experiment with $\beta = 20°$ and $\tau_p = 30\,\text{ms}$. Each line represents a series of single-saltire experiments acquired with the same value of β; τ_p generally decreases going left to right (i.e. a longer pulse means less signal). *Data code:* `5C-190617`.

The results thus obtained are shown in fig. 2.6b. In this plot, both the sensitivities as well as the signal-to-artefact ratios are normalised with respect to the reference double-saltire experiment (black dot at $(1, 1)$, acquired using $\beta = 20°$ and a total $\tau_p = 30\,\text{ms}$, i.e. 15 ms per saltire). An ideal PSE would fall in the top-right corner of this plot. Clearly, as the flip angle increases, the sensitivity increases but at the cost of the purity: this is hardly unsurprising given that the double-saltire experiment has the same property.

While there is no standout candidate which is *clearly* better than the double saltire (i.e. greater

---

[*]As an alternative, the Bruker `sinocal` AU programme was also used to measure the sensitivity of the spectrum; it yielded extremely similar results, so is not shown here. I found the `sinocal` routine to be rather unreproducible as the exact value calculated depends on random noise.

sensitivity as well as purity), the single-saltire pulse with $\beta = 28°$ and $\tau_\mathrm{p} = 30$ ms comes close in performance to the double-saltire experiment. (The use of $\beta = 28°$ here is not coincidental: this flip angle for a single saltire was chosen to (approximately) match the sensitivity of the $\beta = 20°$ double-saltire experiment.) It is generally true that a shorter PSE yields to increased signal, and that usable results were obtained even with a simple 15 ms saltire. However, a shorter PSE also leads to poorer spatiotemporal suppression of artefacts depends and thus poorer purity.

Both the single- and quadruple-saltire cases suffer from the classic dilemma of pure shift NMR: just as in the original PSYCHE, and in the Zangger–Sterk method before it, there is a compromise between sensitivity and purity, and one can only be increased at the cost of the other. Arguably, then, there is not much value in changing the number of saltire pulses as simply varying the *flip angle* of the basic double-saltire already gives the experimentalist a way to balance these competing objectives.

2.4   Direct optimisation of PSYCHE waveform

ureshift__optimisation

The changes to the PSYCHE waveform discussed in the previous section—changing the number of saltire pulses to either 1 or 4—are fairly minor, in that they do not fundamentally alter the form of the PSE. Furthermore, there is an obvious issue in that the quality of the spectrum is not measured in a particularly rigorous manner. Ideally, we would like to have a mathematical way of measuring how well the pure shift sequence has worked. Such a metric would also enable a more automated optimisation process, where a programme is allowed to find the best PSE without any interference from human subjectivity.

In this section, I discuss more radical changes which depart from the tried-and-tested saltire pulse. I also show how two different *cost functions*—functions which determine how 'bad' a spectrum is—can be used to evaluate PSEs. Although the work in this section did not quite yield any groundbreaking results, it provided substantial insights into the nature of pure shift optimisations, which were later used in the context of POISE (chapter 3).

2.4.1   Techniques for pure shift optimisations

hift__optim_techniques

Throughout this chapter (and more generally this thesis), various algorithms are used for *optimisation*: that is, to find the parameters $\mathbf{x} \in \mathbb{R}^n$ which minimise a *cost function $f : \mathbb{R}^n \to \mathbb{R}$*. The theory of optimisation is covered in more detail in § 3.1, so will not be repeated now. Here, I make a distinction only between *derivative-based* and *derivative-free* optimisation algorithms: the former use extra information in the form of ∇f to help locate the optimum, whereas the latter do not, using only the value of $f(\mathbf{x})$. While derivative-based algorithms typically converge to an optimum more quickly, they are unsuitable for problems where the cost function f is noisy.

In this section, the Nelder–Mead (NM) simplex algorithm[62] was used: it is a very popular derivative-free method, and although mathematical convergence is not guaranteed,[63] in practice such cases are rather unlikely to arise. In this section, the implementation of the NM algorithm in the Python `scipy` package was simply used as-is.

The cost function measuring the performance of a pure shift experiment can be measured in one of two ways: either *theoretically*, in that the pure shift experiment is simulated using the density operator formalism, or *experimentally*, in that the experiment is run on a spectrometer. Unfortunately, the simulation of PSYCHE-type PSEs, where a shaped pulse is applied together with a gradient, requires a large amount of time. The pulse itself already has $n \sim 10000$ points, but on top of that, the application of gradients also requires splitting up the sample into multiple slices (often 100 to 1000) such that the evolution of ρ can be simulated in each slice and the results summed up. To make matters worse, H_{pulse} does not commute with H_{grad}, so a propagator

$$U(i, z) = \exp\left[-i(H_{\text{grad}}(z) + H_{\text{free}} + c_x^{(i)} I_x + c_x^{(i)} I_y)(\delta t)\right] \tag{2.29}$$

must be calculated for each pulse point ($1 \leq i \leq n$) in each slice of the sample. As a result, the simulation of PSYCHE spectra in the Spinach package[64] typically requires minutes to hours.*

This in fact makes it faster to experimentally acquire a pure shift spectrum and calculate a cost function based on that. Running an actual pure shift experiment, though, is suboptimal: firstly, the pseudo-2D interferogram method is slow, and secondly, there is no easy way to devise a cost function for the resulting spectrum. Instead, we can use a simple 1D 'pure shift spin echo' (PSSE) sequence, which has the form 90°–τ–PSE–τ–detect (fig. 2.7b). An ideal PSE would lead to complete refocusing of both chemical shifts and J-couplings, and the signal detected after this would simply be the same as in a pulse–acquire experiment (fig. 2.7a). Of course, there is a sensitivity penalty which reflects that of the PSE. On top of that, if the J-refocusing is not perfect, then the multiplets in the spin echo sequence acquire a degree of phase distortion: the delay τ has to be long enough to allow for this, but its exact value is otherwise largely insignificant. These distortions are just about visible in fig. 2.7b.

---

*With highly optimised handwritten code, exploiting the symmetry of the PSYCHE element, the fastest simulation I could do took 16 seconds on a 20-core computer, for a simple 2-spin system. This number increases very rapidly for larger spin systems. It is possible that GPU acceleration could result in substantial speedups, but I have not looked into this sufficiently. Anyway, acquisition of the pure shift spin echo experiment takes only around 5 seconds.

./figures/pureshift/psse.png

Figure 2.7: **(a)** Pulse–acquire experiment and the resulting spectrum. **(b)** Pure shift spin echo experiment and the resulting spectrum. The PSE used was the PSYCHE double saltire, with a flip angle of 25°; the delay τ was 11 ms. The OH peak at 4.1 ppm is lost, most likely due to chemical exchange. *Data code:* `6A-200816`.

Two cost functions were designed and used in this section:

$$f_{\text{phase}} = \text{Var}_i \left[\arctan \left(\frac{S_{\text{re},i}}{|S_{\text{im},i}|} \right) \right] \tag{2.30}$$
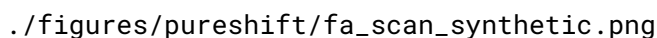
$$f_{\text{diff}} = \left\| \frac{\mathbf{S}_{\text{re}}}{\|\mathbf{S}_{\text{re}}\|} - \frac{\mathbf{T}_{\text{re}}}{\|\mathbf{T}_{\text{re}}\|} \right\|, \tag{2.31}$$

where the PSSE and pulse–acquire 1D spectra are treated as complex-valued vectors $\mathbf{S}$ and $\mathbf{T}$ respectively (for '*spectrum*' and '*target*'). $\mathbf{S}_{\text{re}}$ and $\mathbf{S}_{\text{im}}$ are the real and imaginary parts of the spectrum $\mathbf{S}$, and $S_{\text{re},i}$ is the i-th point of the real part of the spectrum. The operator Var_i represents the variance over all points in the spectrum i, and $\|\mathbf{x}\|$ denotes the 2-norm of the vector $\mathbf{x}$, i.e. $\sqrt{\sum_i x_i^2}$. The implementation of this in Python (listing 2.1) is quite possibly easier to take in.[*]

```
import numpy as np
# assume S and T are complex numpy arrays which have been read in
Sr = np.real(S); Si = np.imag(S); Tr = np.real(T)
f_phase = np.var(np.arctan(Sr / np.abs(Si)))
f_diff = np.linalg.norm((Sr / np.linalg.norm(Sr))
                        - (Tr / np.linalg.norm(Tr)))
```

Listing 2.1: Pure shift cost functions.

---

[*]The use of `np.arctan` (the arctangent), and *not* `np.arctan2` (the argument of a complex number), is intentional. The behaviour shown in fig. 2.8 isn't reproduced with `arctan2`. Of course, this means the name 'phase' is a misnomer; it's not really the phase of anything meaningful.
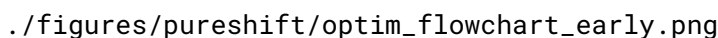
Figure 2.8: Behaviour of the two cost functions, f_{phase} and f_{diff}, on synthetic spectra with various SNRs. Zero phase distortion refers to an in-phase absorption-mode doublet, whereas complete phase distortion refers to an antiphase dispersion-mode doublet. **(a)** The f_{phase} cost function. **(b)** The f_{diff} cost function, measured against a spectrum with no phase distortion and an SNR of 500.

These two cost functions were chosen as they exhibited desirable characteristics on synthetic data (fig. 2.8). In these simulations, the 'target' spectrum was chosen to simply be an in-phase absorption-mode doublet with an SNR of 500. Synthetic data with increasing amounts of phase distortion (i.e. spectra ranging from in-phase absorption, to antiphase dispersion) were generated, and extra Gaussian noise added to create different SNRs. It can be seen that, for data which is well-phased (left edges of the plots), both f_{phase} (fig. 2.8a) and f_{diff} (fig. 2.8b) penalise lower SNRs. Furthermore, both of the cost functions penalise phase distortions in multiplets, since they increase going from left to right. This penalty is stricter for high-SNR spectra, which is also desirable, since it is only in high-SNR spectra that the phase distortions become noticeable.

The cost function f_{diff} is easier to comprehend: it simply scales both the target and PSSE spectra down by their respective intensities, and compares each point to determine whether the peak shapes obtained are similar. Although this seems like it should be agnostic towards signal intensity, this is only true for noiseless spectra. If a (genuine) PSSE spectrum has low SNR, $\|\mathbf{S}_{\mathrm{re}}\|$ will be small, and the noise will be scaled down less than for the target spectrum; this difference in the noise then contributes towards the cost function. On the other hand, a proper rationalisation of why the cost function f_{phase} works is unfortunately not within my capabilities! It was mostly developed by trial-and-error (based on the notion that phase distortions would have something

to do with S_{re} and S_{im}), and I do not have a good explanation of why it works.

Figure 2.9: Flowchart illustrating the steps for optimisation of a pure shift spectrum.

The general optimisation procedure is conceptually simple and largely consists of the loop shown in fig. 2.9: this is essentially a specialised version of the POISE flowchart (CREF TO FIG). The optimisation algorithm is responsible for determining convergence, as well as choosing the new parameters based on previously obtained information; the initial parameters must be supplied by the user.

In practice, it is a technical challenge to implement this loop on a spectrometer as the cost function calculation is performed either in Matlab or Python 3, both of which are not compatible with Bruker's TopSpin software. TopSpin instead provides Jython (Python 2.7) and C programming interfaces;[*] the former is not compatible with Python 3 packages like `numpy` or `scipy`, and the latter is too low-level to be worth implementing numerical algorithms in.[†] Thus, we require a means of *communication* between the spectrometer and the optimisation control programme: this includes a signal from the controlling programme to trigger acquisition on the spectrometer, as well as a signal from the spectrometer that acquisition is done so that the cost function can be calculated. As it turns out, the code used for the optimisations in this section was a very rudimentary and fragile form of that eventually used in POISE (for example, the aforementioned signals were transmitted via the creation and deletion of files). I therefore defer the discussion of this issue to § 3.2, where the more robust POISE interface is explained in detail.

2.4.2 Flip angle optimisation

Having described the rest of the optimisation setup, it remains to choose exactly which parameters are subjected to optimisation. The simplest option is to only optimise one parameter, namely

---

[*]TopSpin 4.1.4 introduced a Python 3 interface which would have made much of this work simpler. Unfortunately, this was not available at the time of this work.

[†]Of course, heavily optimised code in low-level languages such as C and Fortran—or perhaps even Matlab— would run faster. However, speeding up the code has virtually no impact on the optimisation, since its rate is limited by spectrum acquisition. In this situation, it makes far more sense to save *developer time*.

the flip angle of the (double-)saltire pulse. The flip angle dependence of PSYCHE spectra is well-understood, which crucially allows us to evaluate the cost functions outlined above and determine whether they are functioning correctly.

fig:fa_scan_cyclo

Figure 2.10: Behaviour of the two cost functions, f_phase and f_diff, on experimental pure shift spin echo spectra. *Data code:* 5C-190809.

I first sought to measure how the cost functions described above varied with the flip angle. To this end, pure shift spin echo spectra using a *single* saltire as the PSE and various flip angles (from 10° to 50°) were acquired (fig. 2.10). Since both cost functions penalise both low sensitivity and low purity, we might expect that there is be an intermediate value where neither sensitivity nor purity were penalised too much: this would the 'optimum' flip angle. In the event, it was found that only f_diff yielded a useful—albeit shallow—minimum at around 25° (recall that the PSE here is a single saltire, so this corresponds roughly to a 15°–20° double saltire). f_phase, on the other hand, was strictly decreasing within the range of flip angles tested: it is possible that there is an optimum at an even larger flip angle, but this would have been a scientifically unsound conclusion given that the sample was decently concentrated (50 mM).

This naturally raises the question of why f_diff yields an optimum which is sensible. It is tempting to believe that the form of f_diff (eq. (2.31)), which is at first glance quite intuitive, naturally leads to a better result. I argue here, however, that this is mostly down to *coincidence*. This is difficult to explain quantitatively, but in a broader sense, we may imagine that the cost function separately penalises low sensitivity and low purity, i.e. it is of the form

$$f = g(s) + \lambda h(p),\qquad\qquad(2.32)$$ {eq:cost_function_gene

where s and p are respectively the sensitivity and purity, and g and h are some unknown functions which *decrease* with increasing s and p (since we want to minimise the cost function). This

is clearly a simplification, because the plots in fig. 2.8 show that the effects of sensitivity and purity on the cost functions are not additive; however, it is sufficient to make the point here. The parameter λ represents the relative weighting of purity to sensitivity: if λ is large than the purity is more strongly emphasised, and vice versa if λ is small.

When we say that an optimum is 'sensible' or 'sound', this is with respect to the sensitivity/purity balance in the parent pseudo-2D homodecoupled spectrum. In other words, what we *really* seek is a cost function which measures the sensitivity and purity of that spectrum:

$$f' = g'(s') + \lambda'h'(p'), \tag{2.33}$$ {eq:cost_function_gene

where everything is marked with a prime symbol to indicate that it is with respect to the decoupled spectrum, and the parameter λ' is chosen to fit our (admittedly subjective) judgement of the required balance. Now, we may reasonably assume that s and s' are proportional, but p and p' are hardly the same thing: one is manifested in terms of artefact intensity and the other in terms of phase distortions. Furthermore, the λ provided to us by the cost function f may not necessarily be the same as our ideal choice of λ'; let alone the forms of the functions g and h. The fact that eq. (2.32) *happens* to provide the same flip angle optimum as the idealised eq. (2.33) cannot truly be attributed to design!

Of course, just because a cost function works mostly by serendipity does not mean that it cannot be used. So, I ran several *actual* optimisations of the flip angle using the cost function f_{diff}, which reliably converged to optima between 20° and 25° regardless of the initial point chosen. Typically, around 10 spectrum acquisitions were required, corresponding to a time of around 2–4 minutes. This is not surprising in light of our knowledge of f_{diff}; however, it provides us with a proof-of-principle that automated optimisation of NMR parameters is possible. Further optimisations of the PSYCHE flip angle are also discussed in § 3.3.8.

2.4.3   Waveform parameterisation and optimisation

ec:pureshift__chirpopt

Given that a working optimisation setup, including cost functions, had been developed, it was a logical step to then test it out on a more challenging problem: namely, how the waveform used in the PSYCHE PSE could be modified. This goes beyond simply modifying the number of saltires, as was done in § 2.3. There is no real reason why the pulse *must* be an integer number of saltires: in principle it can have *any* shape, although being symmetric about the centre of the pulse would likely still be beneficial in terms of preserving the mechanism of spatiotemporal averaging.

A naive attempt at optimising the pulse would simply involve modifying every pulse point in the double-saltire waveform used in the PSYCHE element. As described in § 1.2, each pulse point consists of a pair of x- and y-amplitudes (c_x, c_y); therefore, for a pulse with n points, we would

have a parameter vector $\mathbf{x} \in \mathbb{R}^{2n}$. Unfortunately, for PSYCHE, n is on the order of 10000, and an optimisation with 20000 points is totally unfeasible.[*]

As a result of this, we must consider other ways of parameterising the waveform. Several approaches to this issue have surfaced in the literature, such as the use of Fourier series,[36,69] Gaussian cascades,[70] or spline interpolation between a subset of pulse points.[71] In this instance, we can use the knowledge that the PSYCHE PSE is composed of saltire pulses to our advantage. Each saltire pulse is a linear combination of two chirps, defined by:

$$\phi(t) = \phi_0 + \pi\tau_{\mathrm{p}}(\Delta F)\left(\frac{t}{\tau_{\mathrm{p}}} - \frac{1}{2}\right)^2 \tag{2.34}$$ {eq:chirp_pulse_phase}

$$c_x(t) = A\cos[\phi(t)] \tag{2.35}$$ {eq:chirp_pulse_cx}

$$c_y(t) = A\sin[\phi(t)] \tag{2.36}$$ {eq:chirp_pulse_cy}

for $t \in [0, \tau_{\mathrm{p}}]$. Here, τ_{p} is the duration of the chirp, A is the amplitude of the chirp (which is time-independent), ϕ_0 the phase of the chirp, and ΔF the bandwidth. (Note that eqs. (2.35) and (2.36) are identical to eq. (1.30).)

Figure 2.11: Plots of the quantities in eqs. (2.34) to (2.36) for a typical chirp pulse ($\tau_{\mathrm{p}} = 30\,\mathrm{ms}$, $\Delta F = 10\,\mathrm{kHz}$). **(a)** $\phi(t)$ wrapped to the range $[0, 2\pi]$. **(b)** The quarter-sine smoothing profile which is later applied to the entire waveform (eq. (2.37)); here $s_{\mathrm{sm}} = 0.2$. **(c)** $c_x(t)/A$ prior to smoothing. **(d)** $c_y(t)/A$ prior to smoothing. The amplitude A is simply a constant which reflects the flip angle of the chirp.[34]

Given these expressions, we see that there are four parameters of the chirp which can be modi-

---

[*]Although problems of this size have been tackled using optimal control theory,[65–68] it is not really feasible to use it in cases where the pulse is applied *together* with a gradient, as is the case in PSYCHE. On top of that, the coupling networks and spin systems of interest are rather more complicated than in typical applications of optimal control.

fied: A, τ_p, ϕ_0, and ΔF. The two chirps which form one saltire pulse simultaneously sweep in opposite directions, which mean that ΔF for one chirp is the negative of the other; however, their parameters are otherwise equal. We may, however, also envision a case where the pulse is constructed from two chirps which are applied at a different point in time, and also cover different bandwidths. This adds two more parameters to each chirp, namely t_0 (the starting time of the pulse) and f_0 (the centre of the pulse bandwidth). Each chirp therefore sweeps from the frequency $f_0 - (\Delta F)/2$ at a time t_0, to the frequency $f_0 + (\Delta F)/2$ at a time $t_0 + \tau_p$. In total, this gives us six parameters per chirp which may be optimised: for a sum of two chirps, there are therefore 12 parameters in total.

Since a sum of two chirps is not necessarily symmetric (with respect to reflection in time), I opted to reflect the waveform about its end, thus doubling the length of the pulse. The entire waveform was then multiplied by a *smoothing function* $f_{sm}(t)$ (fig. 2.11b), which prevents large jumps in RF amplitude at the beginning and end of the pulse. f_{sm} depends on a smoothing parameter s_{sm}, which is typically 0.1–0.2:

$$f_{sm}(t) = \begin{cases} \sin\left(\dfrac{\pi t'}{2 s_{sm}}\right) & 0 \leq t' < s_{sm}; \\ 1 & s_{sm} \leq t' < 1 - s_{sm}; \\ \sin\left[\dfrac{\pi(1 - t')}{2 s_{sm}}\right] & s_{sm} \leq t' \leq 1, \end{cases} \qquad (2.37) \quad \text{\{eq:sming\_function\}}$$

where $t' = t/\tau_p$ (and here τ_p refers to the duration of the *entire* waveform, after reflection.).

The initial point chosen was that corresponding to a single saltire after reflection:

- Chirp 1: $\tau_p = 15\,\text{ms}$; $\Delta F = -5\,\text{kHz}$; $\phi_0 = 0$; $A = 36\,\text{Hz}$; $t_0 = 0$; $f_0 = 5\,\text{kHz}$;

- Chirp 2: $\tau_p = 15\,\text{ms}$; $\Delta F = 5\,\text{kHz}$; $\phi_0 = 0$; $A = 36\,\text{Hz}$; $t_0 = 0$; $f_0 = -5\,\text{kHz}$;

Collectively, these two pulses sum up to become the *first half* of a single saltire with bandwidth 10 kHz. After reflection, the total duration of the saltire is 30 ms, and the amplitude is 72 Hz, which corresponds to a flip angle of approximately 32°.

Using this setup, several optimisations of the 12 parameters above were conducted. It was quickly noticed that, although f_{diff} was a better cost function than f_{phase}, this led to spurious optima being located, such as the one in fig. 2.12c. Although the initial point (a saltire) yielded a much larger SNR (fig. 2.12b), the value of f_{diff} was still larger than for this false optimum. The reason for this is almost certainly because the PSE distorts the relative intensity of the strong singlets in the spectrum, notably the C_6D_5H peak at 7.15 ppm, and the N-methyl groups between 2.5 and 3.8 ppm).

Of course, singlets are completely unimportant when devising a pure shift experiment. Unfortu-

Figure 2.12: (**a**) Target spectrum (pulse–acquire). (**b**) PSSE spectrum obtained using the initial guess (a saltire). (**c**) PSSE spectrum obtained with a spurious optimum point. The grey dotted box shows the restricted region over which f_{diff} was subsequently applied to, yielding a formally different cost function f'_{diff}; this yielded more sensible results where the cost function for the saltire pulse was smaller. *Data code:* 6C-190823.

nately, since singlets are typically more intense than the rest of the spectrum, they also contribute disproportionately to the cost function. A simple and effective way to circumvent this is to restrict the part of the spectrum being evaluated. In this case, I chose to use the H^{α} region between 4.72 and 5.94 ppm (grey dotted box in fig. 2.12). This yields a formally different cost function, which I label as f'_{diff}. With this, much more logical behaviour was observed: in particular, the saltire pulse performed better than the spurious optimum previously found.

While this new cost function could be successfully used to run optimisations, unfortunately most of these simply failed to find anything performing better than the original saltire pulse. On the rare occasion where something 'better' was found (as judged by the new cost function f'_{diff}), these were simply *fairly close* to that of a saltire, and the corresponding decreases in the cost function extremely small—suggesting that the 'better' result may simply just have been due to noise in the cost function. Nevertheless, these new 'optima' *did* function as perfectly serviceable PSEs: for example, fig. 2.13 compares a TSE-PSYCHE spectrum obtained with an optimised pulse to one obtained with the single saltire pulse. There is virtually no difference. This is a meaningful result, as it demonstrates that f'_{diff} is actually an accurate metric to determine the quality of a PSE (it is noisy, but this is inevitable in an experimentally measured cost function). Unfortunately, although the pulse shape is shown in fig. 2.13b, the exact parameters which generated this pulse shape have been lost to time.

Figure 2.13: **(a)** *x*- and *y*-coefficients of the initial saltire pulse (as a fraction of the maximum amplitude *A*). **(b)** *x*- and *y*-coefficients of the 'optimised' pulse (as a fraction of the maximum amplitude *A*). **(c)** TSE-PSYCHE spectrum obtained with the initial guess (a saltire pulse). **(d)** TSE-PSYCHE spectrum obtained with the 'optimised' pulse, with virtually equivalent performance. *Data code:* `6C-190831`.

In conclusion, while not quite as intractable as 20000 parameters, optimising a 12-parameter function clearly still proves to be a challenge. Although the cost functions described here do work, they are generally quite 'flat', in that they do not discriminate very sharply between 'good' and 'bad' spectra. Combined with the fact that the cost function is noisy, this makes experimental optimisation of the waveform an uphill task. Nevertheless, much of the knowledge (and code) in this section was later used in the development of POISE. In the next two sections in this chapter, I will move on from PSYCHE and discuss completely different methods of obtaining pure shift spectra using only hard pulses in the PSE.

2.5 Time-reversal method

sec:pureshift__timerev

In § 2.2.2, the time-reversal element was briefly introduced; I now provide a more theoretical description of it. As stated in that section, this merely consists of a hard pulse with a flip angle β, and using eqs. (2.25) to (2.27) we first analyse its potential use as a JRE. Recall that an ideal JRE should invert passive spins only, not active spins: thus, we seek a transformation of the form

$I_{1+}I_{2\alpha} \to I_{1+}I_{2\beta}$. However, the hard pulse does this:

$$I_{1+}I_{2\alpha} \to \underbrace{c^2 s^2 I_{1+}I_{2\beta}}_{\text{term 1}} + \underbrace{c^4 I_{1+}I_{2\alpha}}_{\text{term 2}} - \underbrace{\frac{S^2}{4}I_{1\alpha}I_{2+} + \frac{S^2}{4}I_{1\beta}I_{2+}}_{\text{terms 3 and 4}}, \qquad (2.38)$$ `{eq:beta_pulse_not_jre`

where (as before) $S = \sin\beta$, $s = \sin(\beta/2)$, and $c = \cos(\beta/2)$. As in the analysis for PSYCHE (§ 2.2.3), term 1 here represents the desired signal; term 2 recoupling artefacts; and terms 3 and 4 COSY-type artefacts arising from coherence transfer. Terms with difference coherence orders have been neglected as they can be removed using CTP gradients.* Unlike in the PSYCHE analysis, however, we have not neglected any other terms of smaller order in *s*.

Since the desired and undesired terms have different coefficients, it is possible to *fully cancel out the recoupling artefacts* by recording (in this case) two different spectra with different values of β and performing an appropriate linear combination:

$$I_{1+}I_{2\alpha} \xrightarrow{\beta=0°} I_{1+}I_{2\alpha} \qquad (2.39)$$ `{eq:timereversal_twosp`

$$I_{1+}I_{2\alpha} \xrightarrow{\beta=90°} \frac{1}{4}I_{1+}I_{2\beta} + \frac{1}{4}I_{1+}I_{2\alpha} - \frac{1}{4}I_{1\alpha}I_{2+} + \frac{1}{4}I_{1\beta}I_{2+} \qquad (2.40)$$ `{eq:timereversal_twosp`

If we take eq. (2.40) minus 1/4 of eq. (2.39), the recoupling artefacts (arising from the $I_{1+}I_{2\alpha}$ term) are fully removed. In general, for an *N*-spin system, there are several different 'types' of recoupling artefacts where different numbers of passive spins (between 1 and $N-1$) are not inverted. Each of these pathways will have different coefficients, as each spin that is flipped contributes s^2, whereas each spin that is not flipped contributes c^2. Suppressing all of these requires the acquisition and summation of *N* spectra with different flip angles and appropriate weights.

Before we go on further, notice even in the two-spin system that the COSY-type artefacts are *not* suppressed! In the original work,[13] this time-reversal element was used in the middle of the t_1 period in a NOESY experiment. The coherence transfer peaks were not deemed to be problematic in this context: they gave rise to artefacts which had F_1 frequencies of $(\Omega_1 + \Omega_2)/2$, but different phase properties to genuine crosspeaks, allowing them to be easily identified. Of course, these artefacts are not acceptable in an actual pure shift spectrum.

To remove these peaks, I adopted the strategy first reported by Thrippleton *et al.* for suppression of COSY-type transfer pathways in 2DJ spectra.[4] In a 2DJ experiment, the central 180° pulse should in principle not cause coherence transfers between different spins; however, in strongly coupled systems this can happen. The solution was to bracket the 180° pulse, as well as half of the t_1 period, with a pair of opposing chirps and gradients. The same idea was later used in the TSE-

---

*In the original work[13] which predated widespread use of field gradients, other terms were removed through phase cycling, which is essentially equivalent.

PSYCHE experiment[11] to (further) suppress strong coupling responses in the parent PSYCHE experiment. This works because the unwanted CTPs have coherences on different spins during each of the two chirps; consequently, the coherences are inverted at different times by the chirp pulses, and are ultimately dephased by gradients. The resulting time-reversal experiment was thus simply the same as the parent TSE-PSYCHE experiment, except that the central PSYCHE element was replaced by a β hard pulse (fig. 2.14).
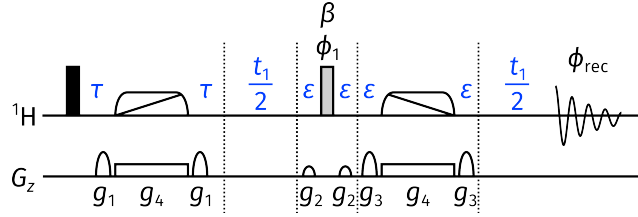
Figure 2.14: Time-reversal pure shift pulse sequence. The flip angle β is varied as described in eq. (2.41). Pulse phases are: $\phi_1 = (x, y, -x, -y)$; $\phi_{\text{rec}} = (x, -x, x, -x)$. The delay τ is set to $1/(4 \cdot T_{\text{chunk}})$, and allows for J-coupling to be refocused in the middle of the chunk. Gradient amplitudes are $(g_1, g_2, g_3, g_4) = (35\%, 49\%, 77\%, 1\%)$ (note that, in principle, g_4 should be calibrated according to the bandwidth of the chirp used).

In such a sequence, the JRE is in fact not only the β pulse itself but also the second chirp (which effectively acts as a 180° pulse). So, the β pulse here fulfils the role of a PSE, not a JRE: this means that different flip angles and weights must be chosen. As before, for a system containing N mutually coupled spins, N different experiments must be acquired:

$$\beta_j = \frac{j\pi}{N} \tag{2.41}$$

$$W_j = \frac{N}{8} \cdot \frac{(-1)^j}{\sin^2(\beta_j/2)} \tag{2.42}$$

where $j = 1, 2, \ldots, N$. For the sake of completeness, the values of β and W for a JRE are given here as well:

$$\beta_k = \frac{k\pi}{N} \tag{2.43}$$

$$W_k = \frac{N}{8} \cdot \frac{(-1)^{k+N}}{\cos^2(\beta_j/2)} \tag{2.44}$$

for $k = 0, 1, \ldots, N - 1$. The derivation of these expressions is discussed more thoroughly in a paper by Griesinger et al.[5] Note that the weights for the PSE correspond to that used for the ECOSY experiment, and the weights for the JRE correspond to that used for the complementary ECOSY experiment.

In practice, $N = 5$ is likely to cover most realistic spin systems: this means that we have $\beta = \{36°, 72°, 108°, 144°, 180°\}$ and $W = \{-6.545, 1.809, -0.955, 0.691, -0.625\}$. Figure 2.15 shows

insets of the five subspectra acquired with the above values of β and scaled by their respective weights. All of the spectra have been phased using the automatic phase corrections determined for the weighted sum (i.e. the pure shift spectrum).
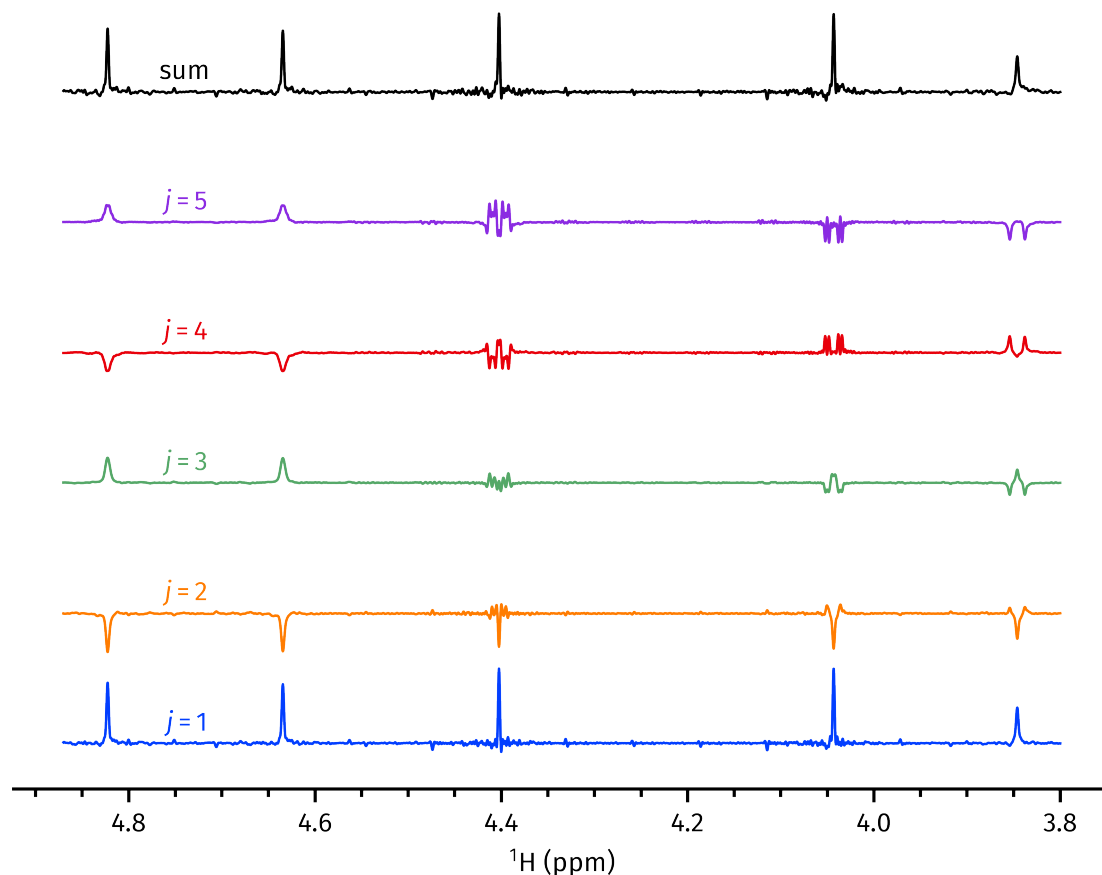

Figure 2.15: Insets of weighted time-reversal subspectra (with $j = 1$ through 5), as well as their sum (the pure shift spectrum). *Data code:* `7A-201020`.

Although the experiment seems to work, in that the sum *is* indeed a pure shift spectrum, the fact that it is obtained through summation of N different experiments brings some immediate drawbacks. Firstly, the minimum duration of the experiment is lengthened by a factor of N: this is essentially the same as an N-step phase cycle. However, and perhaps more importantly in the context of *pure shift* NMR, the artefacts surrounding the main peaks are not perfectly cancelled through the process of summation. As a result, random distortions are observed around the desired peaks in the pure shift spectrum: this is noticeable in the 4.04 ppm peak in fig. 2.15, and is even worse for intense signals.

In terms of sensitivity, the time-reversal spectrum is not particularly exceptional, either. Each of the five subspectra above were acquired with 2 scans; when compared against a typical TSE-PSYCHE experiment acquired with only 4 scans (i.e. 40% of the experiment duration), the TSE-PSYCHE experiment had comparable or perhaps even slightly better SNR (fig. 2.16). The likely reason for this is because in the time-reversal experiment, signal is actually being *cancelled out*

through the process of summation, as is quite clearly shown in fig. 2.15. In principle, the sensitivity of the time-reversal experiment could be optimised by acquiring the more heavily weighted spectra with more scans. However, I consider this unlikely to make a substantial difference to the conclusions drawn here. The idea of optimising the weights to better suppress artefacts was also briefly considered: however, given that eq. (2.42) already yields *theoretically* complete suppression, it was deemed unlikely that anything substantially better could be obtained.

(a)   **(b)**
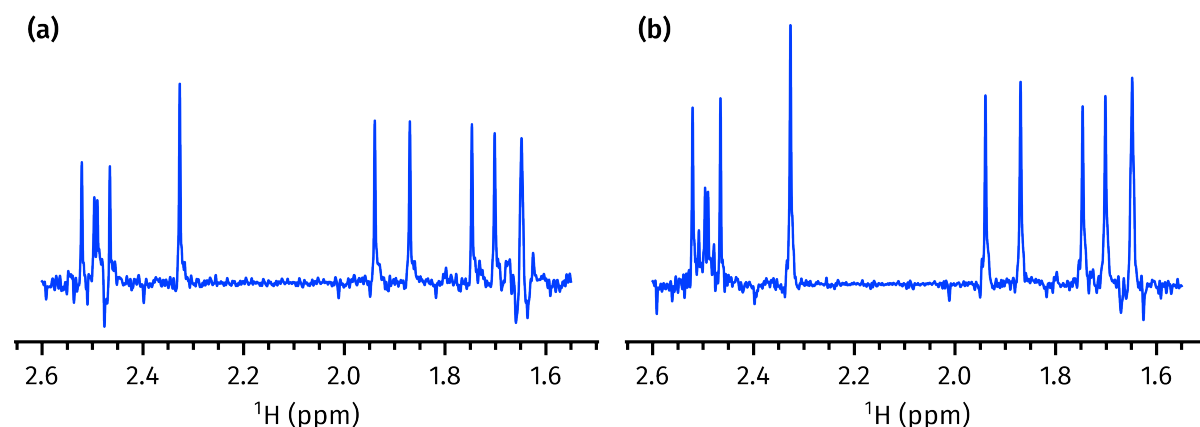
Figure 2.16: Comparison of time-reversal and TSE-PSYCHE sensitivity. **(a)** Time-reversal pure shift spectrum (the same as the sum in fig. 2.15) acquired with 2 scans for each sub-spectrum. **(b)** TSE-PSYCHE (double saltire, flip angle 15°) acquired with 4 scans. The spectra have been scaled so that their noise levels are similar: the signal intensity is comparable, or perhaps slightly better in the TSE-PSYCHE. *Data code:* `7A-201020`.

Interestingly, fig. 2.15 suggests that the $j = 1$ spectrum *on its own* already provides as good a result as the summed pure shift spectrum. This is not surprising, as the use of a hard pulse as the PSE yields a conceptually very similar result to PSYCHE, in that the signal to artefact ratio depends on $\tan^2(\beta/2)$ (of course, the COSY-type artefacts must still be suppressed through the TSE scheme). This suggests that the 'TSE' time-reversal experiment in fig. 2.14 is a viable pure shift sequence without requiring the summation of multiple subspectra—albeit one which does not have any significant advantage over PSYCHE.

2.6   'Discrete PSYCHE'

The last pure shift method in this chapter is completely original, and represents perhaps the most fruitful attempt so far at optimising pure shift experiments. It relies on what is essentially a 'temporal discretisation' of the PSYCHE waveform and gradient combination: instead of applying a shaped pulse and a gradient simultaneously, hard pulses and gradients are interleaved in the PSE (fig. 2.17).

For this reason I have dubbed this experiment the 'discrete PSYCHE', or dPSYCHE for short.
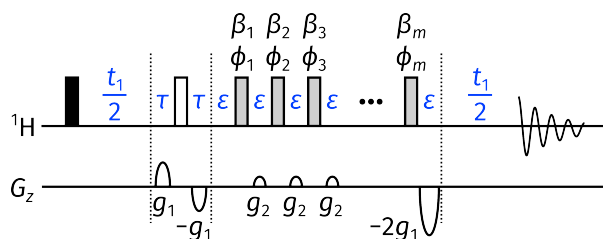
Figure 2.17: dPSYCHE pulse sequence. Gradient amplitudes are $(g_1, g_2) = (35\%, 41\%)$; the gradients in the PSE g_2 are applied with a duration of $500\,\mu s$. The hard pulses in the PSE are applied with an RF amplitude of $18\,kHz$. The delay τ is set to $1/(4 \cdot T_{\text{chunk}})$, and allows for J-coupling to be refocused in the middle of the chunk.

There are two major reasons why this is more amenable towards optimisation than many of the previous experiments:

1. Pulses and gradients are no longer applied simultaneously, which makes simulation of the experiment *extremely* fast compared to the original PSYCHE. This opens up the possibility of entirely theoretical optimisations, as the noise can be completely eliminated from the cost function.

2. There are much fewer 'pulse points' than in the original PSYCHE: effectively, the phase and flip angle of every hard pulse has to be optimised, leading to $2m$ parameters. Even for $m \sim 10$, this is quite tractable if the optimisation is noiseless.

One downside of this is that it is difficult, or perhaps even impossible, to explain how the PSE works.* For a symmetric PSE where $\beta_1 = \beta_m$ (and so on) it is probably possible to reuse an explanation based on PSYCHE-style CTP selection, but this is clearly inapplicable if the flip angles and phases are scrambled.

2.6.1   Speeding up dPSYCHE simulations

To begin, I first explain how the exact simulation of dPSYCHE experiments can be greatly accelerated through efficient propagator calculations. Although Spinach[64] is the leading simulation package for NMR, and covers an extremely impressive range of experiments, this generality also prevents it from providing optimal performance for any single experiment. As it turns out, handwritten, specialised Matlab code can outperform Spinach by orders of magnitude.

The NMR simulations developed here simply use the density operator formalism in Hilbert space, as outlined in § 1.3. The Zeeman basis is used, and non-unitary transformations such as relaxation are neglected. Propagation under the Liouville–von Neumann equation (eq. (1.38)) requires

---

*Of course, we could simulate it and say that it works because the maths says it does; but that isn't very illuminating. The optimisations done in this thesis are somewhat like a scaled-down version of machine learning, in that they produce better results at the cost of interpretability.

matrix exponentials $\exp(-iHt)$. For an $N \times N$ matrix, the matrix exponential requires $O(N^3)$ time to calculate (and for a system containing p spins, we have $N = 2^p$); it is often this which is the bottleneck in NMR simulations. Minimising the number of matrix exponentials, and/or their computational cost, is the key to achieving speedups, as will be shown in the following text.[*]

Generally, the accurate simulation of gradient-enhanced experiments requires the sample to be divided up into n discrete slices, each simulated with a different H_{grad}.[†] Thus, a very naive implementation of the dPSYCHE experiment would require mn matrix exponentials, one per pulse per slice. The overall structure of the code would resemble listing 2.2. (Strictly speaking, the above code is wrong because the final gradient should have strength `-2G` and not `G`, but that is a minor detail.)

```matlab
% loop over slices
for slce=1:n
    H_grad = I_z * G * z(slce);

    % loop over pulse points
    for pulse_point=1:m
        H_pulse = (c_x(pulse_point) * I_x) + (c_y(pulse_point) * I_y);

        % calculate propagators; m*n total matrix exponentials
        U_pulse = expm(-1j * (H_free + H_pulse) * t_pulse);
        U_grad = expm(-1j * (H_free + H_grad) * t_grad);

        rho = U_grad * U_pulse * rho * U_pulse' * U_grad';
    end
end
```

Listing 2.2: Rough structure of a naive dPSYCHE implementation. Note that I use the variable name `slce` as `slice` is an existing builtin Matlab function.

It is not difficult to come up with a more sensible approach which cuts this down by a factor of m: since the pulses are not applied together with the gradients, the pulse propagators U_{pulse} can be pre-calculated outside of the loop. Furthermore, all of the gradients within the PSE are the same, so U_{grad} can be moved out of the inner loop (listing 2.3).

---

[*]Note that in my simulations, I simply used the builtin `expm` Matlab function, which implements the matrix exponential using a combination of the scaling-and-squaring method and Padé approximation.[72] This is in contrast to Spinach, which uses a scaled-and-squared Taylor series (according to the `propagator.m` file, the other methods supposedly did not 'live up to their marketing'). An in-depth discussion of matrix exponential methods is outside the scope of this thesis, but can be found in a classic paper by Moler and Van Loan.[73]

[†]In simple cases this can be avoided by simply removing all terms with the wrong coherence orders as we know they will be dephased (eq. (1.110)), but this is too naive an approach for pure shift techniques.

```matlab
% precalculate pulse propagators; m total matrix exponentials
for pulse_point=1:m
    H_pulse = (c_x(pulse_point) * I_x) + (c_y(pulse_point) * I_y);
    U_pulse(m) = expm(-1j * (H_free + H_pulse) * t_pulse);
end

% loop over slices
for slce=1:n
    % calculate gradient propagators; n total matrix exponentials
    H_grad = I_z * G * z(slce);
    U_grad = expm(-1j * (H_free + H_grad) * t_grad);

    % loop over pulse points
    for point=1:m
        rho = U_pulse(m) * rho * U_pulse(m)';
        rho = U_grad * rho * U_grad';
    end
end
```

Listing 2.3: Rough structure of a slightly faster implementation of dPSYCHE.

Spinach, which is designed to be general, has no idea that these optimisations are possible, so is naturally rather slower. However, even this is relatively inefficient. Notice that the two components of the gradient propagator, H_{free} and H_{grad}, actually commute with one another (even in the strong coupling case). (Showing that $[H_{\text{cs}}, H_{\text{grad}}] = 0$ is trivial; the algebra to show that $[H_{\text{J}}, H_{\text{grad}}] = 0$ is slightly more tedious, but does not require any special techniques.) Thus, we can write:

$$\exp[-\mathrm{i}(H_{\text{free}} + H_{\text{grad}})t] = \exp(-\mathrm{i}H_{\text{free}}t)\exp(-\mathrm{i}H_{\text{grad}}t) \tag{2.45}$$ {eq:split_propagator}

(in general, for matrices A and B, $\exp(A + B) = \exp(A)\exp(B)$ if and only if $[A, B] = 0$). This on its own does not reduce the number of matrix exponentials required, but notice now that H_{grad} is a sum of I_{iz} terms and is therefore *diagonal* in the Zeeman basis. The exponential of a diagonal matrix is trivial to calculate, as we simply need to exponentiate the diagonal *elements:*

$$\exp\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = \begin{pmatrix} \exp(\lambda_1) & 0 & \dots & 0 \\ 0 & \exp(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \exp(\lambda_n) \end{pmatrix}. \tag{2.46}$$ {eq:expm_diagonal}

This can be done in $O(n)$ time by using `diag(exp(diag(M)))` instead of `expm(M)` (the `diag` Matlab function converts a diagonal matrix to a vector of its diagonal entries, and vice versa). So, our code now uses only $m + 1$ total matrix exponentials (listing 2.4). At this point, our matrix exponentials have almost been eliminated and the largest remaining bottleneck is almost certainly

the matrix *multiplications* required for the propagation. We can cut this down by a factor of two simply by calculating the overall propagator

$$U_{\text{total}} = U_n \cdots U_2 U_1 \tag{2.47}$$ {eq:overall_propagator}

and then performing the propagation only at the very end:

$$\rho = U_{\text{total}} \rho_0 U_{\text{total}}^{\dagger} \tag{2.48}$$ {eq:overall_propagatio}

instead of performing every individual step $\rho \rightarrow U_1 \rho_0 U_1^{\dagger}$. The final optimised code therefore resembles that in listing 2.4.

```
% precalculate propagator due to free evolution during gradient
% only 1 matrix exponential
U_free = expm(-1i * H_free * t_grad);

% precalculate pulse propagators; m total matrix exponentials
for pulse_point=1:m
    H_pulse = (c_x(pulse_point) * I_x) + (c_y(pulse_point) * I_y)
    U_pulse(m) = expm(-1j * (H_free + H_pulse) * t_pulse)
end

% loop over slices
for slce=1:n
    % initialise propagator for this slice
    U_slce = eye(2 ^ p);

    % calculate gradient propagators; no matrix exponentials required
    H_grad = I_z * G * z(slce);
    U_grad = U_free * diag(exp(diag(-1j * H_grad * t_grad)));

    % loop over pulse points
    for point=1:m
        U_slce = U_pulse(m) * U_slce;
        U_slce = U_grad * U_slce;
    end

    % perform propagation only at the end
    rho = U_slce * rho * U_slce';
end
```

Listing 2.4: Rough structure of a fast dPSYCHE implementation.

lst:dpsyche_fast2

The performance of this handwritten code as compared to Spinach is summarised in table 2.1. In all cases, the spectra produced by the two methods were entirely equivalent. It should be noted that the handwritten code does *not* even utilise CPU parallelisation, whereas Spinach

does. I investigated the possibility of parallelising the loop over slices (replacing the outer `for` in listing 2.4 with `parfor`): however, this in fact made the code *slower*, presumably due to overhead. This is a good thing: it means that `parfor` can be used in an external loop, e.g. for the parallel simulation of the dPSYCHE experiment on different spin systems.

Number of spins	Number of couplings	Execution time (s)	
		Spinach	Handwritten
1	0	3.33	0.35
2	0	4.59	0.32
	1	6.22	0.32
3	0	9.90	0.43
	1	12.95	0.45
	2	31.86	0.47
	3	35.01	0.48
4	0	30.59	1.02
	1	38.63	0.99
	2	44.61	1.04
	3	365.04	1.52
	4	446.03	1.57
	5	521.72	1.57
	6	588.31	1.69

Table 2.1: Comparison of wall-clock times for dPSYCHE simulations. The dPSYCHE sequence used contained 15 pulses, each applied with a flip angle of 15° and a phase of 0°. Spin systems were generated pseudo-randomly. All timings were measured on a 2019 MacBook Pro with a 2.6 GHz 6-core Intel i7 CPU; no attempts were made to shut other running programmes.

2.6.2 Optimisations and experimental evaluation

As previously discussed, the fact that the dPSYCHE experiment can be very quickly simulated opens up the potential for entirely computational optimisation of the pulse sequence. For any arbitrary spin system, it is trivial to remove all the couplings *in silico* and simulate a pulse–acquire spectrum: this gives us a theoretically perfect pure shift spectrum. The simulated dPSYCHE experiment (on a system with couplings) can then be compared against this. The entire process is repeated using n different spin systems, and the cost function is defined as:

$$f_{\text{diff},2} = \frac{1}{n} \sum_n \left\| \frac{\mathbf{S}_{\text{re}} - c\mathbf{T}_{\text{re}}}{c\|\mathbf{T}_{\text{re}}\|} \right\|^2 \tag{2.49}$$

where **S** is the dPSYCHE spectrum, and **T** is the target spectrum. c is a factor to be discussed later, and for now I treat it as 1. This cost function appears superficially very similar to f_{diff} (discussed in § 2.4.1), and is based on the same idea that we want the spectra **S** and **T** to match one another, but there are several points of note:

- The spectrum $\mathbf{S}_{\text{re}}$ is not scaled down by its norm. This means that the sensitivity penalty no longer comes from the difference in the noise (as was previously the case), but rather directly from the difference in peak intensity. Since simulated spectra are noiseless, the original f_{diff} would not work here.

- The division of the entire cost function by $\|\mathbf{T}_{\text{re}}\|$ is not important if only one spin system is being simulated as it is simply a constant factor. However, if more than one spin system is being simulated, $\|\mathbf{T}_{\text{re}}\|$ differs from system to system and this factor helps to essentially normalise the contributions from each spin system.

- The norm in the cost function here is squared. Again, this makes no difference to the optimum if only one spin system is being investigated, because x^2 is strictly increasing for $x > 0$.* However, for multiple spin systems it makes sense to square the norm, as the largest deviations will be penalised more greatly: this means that a pure shift spectrum which works reasonably well across a wide range of spin systems will be prioritised over one which works perfectly well for some and fails badly for others.

I began by first checking how many t_1 increments (i.e. chunks) were required in the simulation to obtain reliable cost function values. If too few chunks are simulated, the resulting pure shift spectrum will have truncation artefacts, which are likely to mask artefacts from unwanted CTPs. The value of $f_{\text{diff},2}$ was thus tested with a wide variety of randomly chosen phases and angles, with the number of chunks set to 4, 8, and 16 (fig. 2.18).

The 16-chunk and 8-chunk $f_{\text{diff},2}$ (fig. 2.18a) do in fact line up quite well. Notably, as the inset shows, they both agree on the 'best' candidate (note that this is not necessarily anywhere near perfect, since it was randomly generated). The 4-chunk $f_{\text{diff},2}$ also has the correct overall behaviour (fig. 2.18b). However, its ranking of the 'best' candidates is not very accurate: the 4-chunk cost function rates the red dot in the inset as the optimum, but that is only the 13th-best candidate when using the 16-chunk cost function. Ultimately, I decided to use the 4-chunk cost function for 'quick and dirty' optimisations, where only an approximate optimum was required. However, for anything requiring more accuracy, the 8-chunk cost function was used.

Optimisations were next carried out with an (arbitrarily chosen) setting of $m = 9$ (i.e. nine hard pulses in the PSE). The derivative-based Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm was used to carry out the optimisation: this is a popular line search algorithm which

---

*It may affect the rate of convergence, but this is not something I tested.
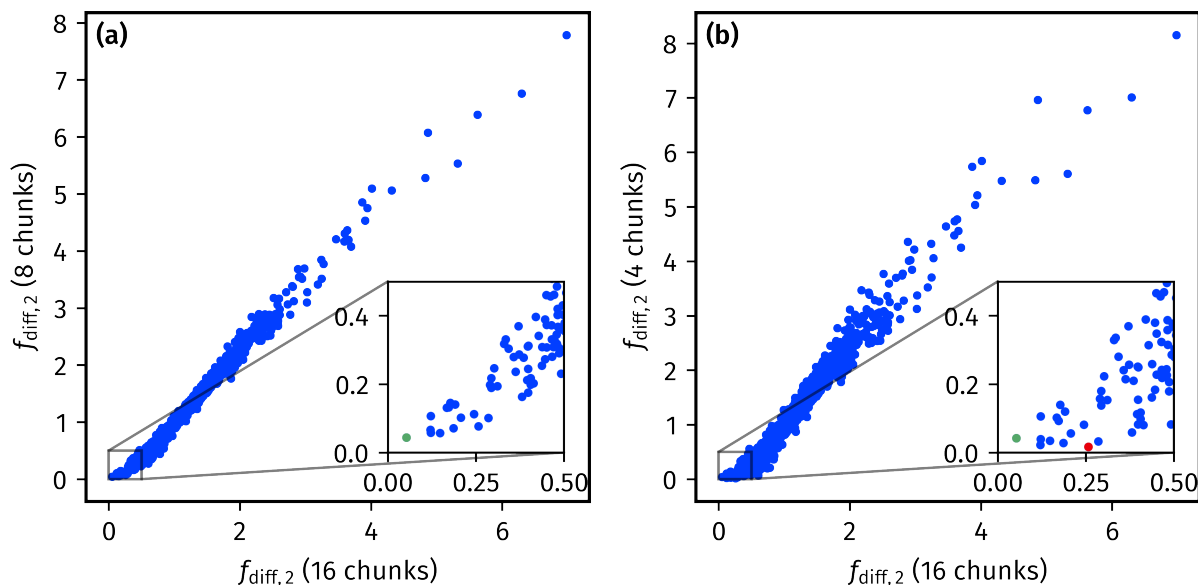
Figure 2.18: Comparison of $f_{\text{diff},2}$ values when simulated with different numbers of chunks. 16 chunks is assumed to be the 'gold standard'. **(a)** Correlation between 16-chunk and 8-chunk cost functions. The 'optimum' identified by both cost functions is plotted in green in the inset. **(b)** Correlation between 16-chunk and 4-chunk cost functions. The 'optima' identified by the 16- and 4-chunk cost functions are plotted in green and red respectively in the inset.

uses an approximate Hessian to calculate the search direction at each iteration.[74,75] No lower or upper bounds were placed on the flip angles or phases: phases can obviously simply be wrapped to the region $[0, 2\pi)$, and in the simulations, the hard pulses were modelled as being instantaneous rotations, so their flip angles can also just be wrapped to $[0, 2\pi)$. (This would not be completely valid if realistic, finite pulses were used, since changing the flip angle would also change their duration.) This first optimisation yielded the following optimised parameters for the nine pulses:

$$\{\beta_i\} = \{118.6972°, 29.8400°, 107.4850°, 190.4788°,$$
$$138.4710°, 144.5939°, 18.9674°, 73.8900°, 130.6071°\} \qquad (2.50) \quad \text{\{eq:dpsyche\_opt\_nosens}}$$

$$\{\phi_i\} = \{144.5641°, 173.3596°, 38.9878°, 146.3121°,$$
$$127.7346°, 7.5104°, 36.9805°, 110.2791°, 182.3894°\}. \qquad (2.51) \quad \text{\{eq:dpsyche\_opt\_nosens}}$$

The fact that none of the flip angles exceeded 360° here provides some justification for not using bounds in the optimisation.

To evaluate the quality of the decoupling on a real sample, the dPSYCHE experiment was experimentally acquired. fig. 2.19 compares the results against a PSYCHE experiment. The dPSYCHE experiment clearly has much greater sensitivity than the PSYCHE experiment; however, the decoupling quality is extremely poor. The reason for this is likely because the optimisation is
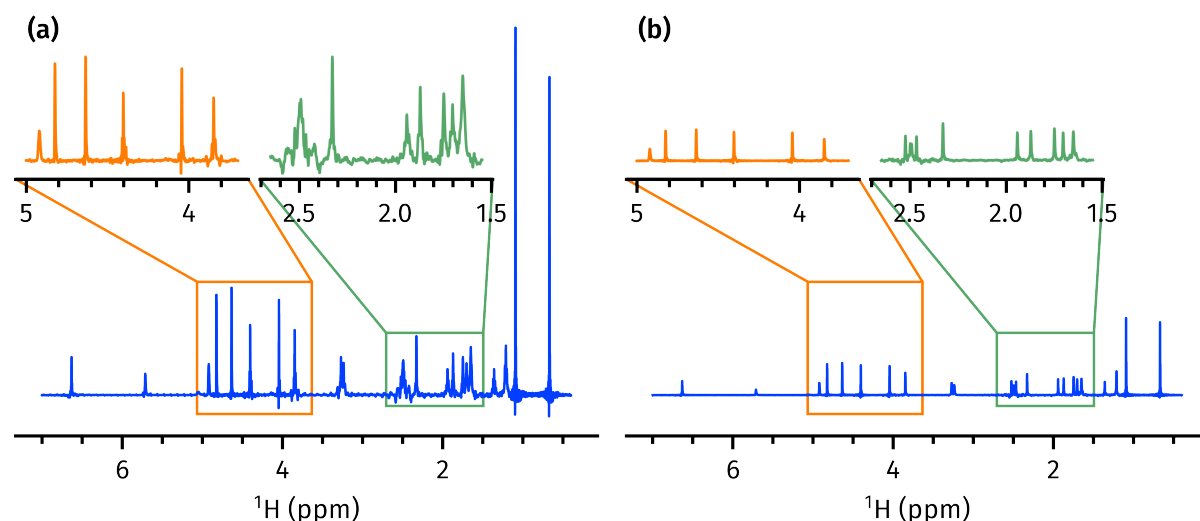
Figure 2.19: **(a)** dPSYCHE experiment acquired using the flip angles and phases in eqs. (2.50) and (2.51). The β hard pulses in the PSE were applied with an RF amplitude of **(b)** (Double-saltire) PSYCHE experiment acquired with $\beta = 20°$. *Data code:* 6A-211027.

prioritising sensitivity too highly over purity, and ultimately stems from the factor c in eq. (2.49) which we have ignored until now. Since $c = 1$, the optimisation is essentially guiding the dPSYCHE experiment towards having *exactly the same sensitivity as a pulse–acquire experiment.* Although this might be a nice idea in principle, it is not physically sensible: no pure shift method has sensitivity which is even close to that of a pulse–acquire experiment. It makes more sense to scale down the target spectrum by introducing a factor of $c < 1$ into the cost function (eq. (2.49)); the parameter c therefore represents the 'target sensitivity'. By changing the parameter c, we can control whether the optimisation emphasises sensitivity or purity more. (Note that this option was *not* available to us in the experimental PSSE-based optimisations in § 2.4.)

A series of new optimisations were therefore run, with c ranging from 0.2 to 1 (TABLE). The 4-chunk cost function was used, and the maximum number of function evaluations capped at 5000. Each optimisation was run 10 times with a different starting seed, and the best of the 10 results chosen for further evaluation (table 2.2). The resulting spectra (fig. 2.20) show that adjusting c has the desired effect: larger values lead to greater sensitivity and lower purity, and vice versa.

In this spectrum of andrographolide, the 3.5 ppm to 5 ppm region is an 'easier' region to decouple: there are fewer couplings here and all peaks are firmly within the weak coupling regime. The 1.5 ppm to 2.8 ppm region is 'more difficult' to decouple, especially the peaks at 2.5 ppm, which are mutually strongly coupled. The dPSYCHE spectrum with $c = 0.4$ appeared to be a good compromise between purity and sensitivity: Furthermore, its sensitivity was higher than that of the original PSYCHE experiment (when conducted with an 20° double saltire) Furthermore, it
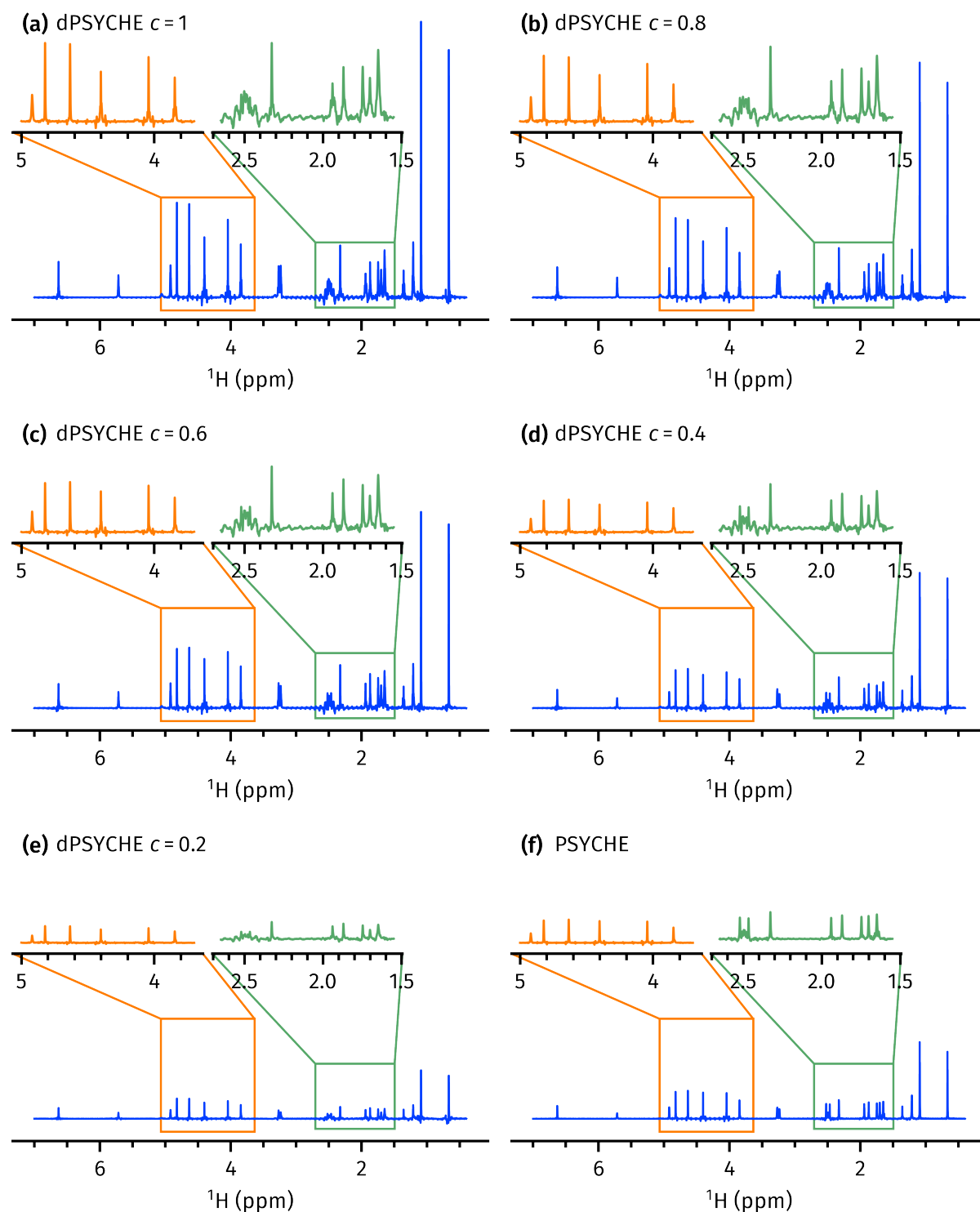
Figure 2.20: **(a)** dPSYCHE, $c = 1$. **(b)** dPSYCHE, $c = 0.8$. **(c)** dPSYCHE, $c = 0.6$. **(d)** dPSYCHE, $c = 0.4$. **(e)** dPSYCHE, $c = 0.2$. **(f)** Double-saltire PSYCHE with $\beta = 20°$. All spectra and insets are plotted to scale to allow for sensitivity comparisons. *Data code:* 6A-211231.

c	Flip angles (°)	Phases (°)	$f_{\text{diff},2}$
1	59.8813, 96.0748, 111.5862, 118.7285, 144.701, 176.9773, 29.2866, 40.0658, 63.8237	355.773, 81.741, 99.8752, 81.6675, 337.0056, 150.7004, 105.9666, 317.9979, 27.7897	0.0404
0.8	81.9461, 65.311, 90.3488, 106.6388, 73.3462, 196.7857, 51.791, 34.8188, 53.9506	329.8571, 60.3564, 137.3929, 68.5286, 340.74, 126.29, 64.4601, 145.4714, 27.3019	0.0321
0.6	53.1313, 82.2547, 88.6093, 161.9291, 83.462, 161.3548, 14.2976, 66.2199, 53.9709	347.2835, 59.5518, 68.072, 113.5358, 34.5537, 156.191, 268.7641, 302.7427, 16.29	0.0250
0.4	77.2998, 127.7274, 87.6663, 104.6371, 99.6474, 168.8171, 8.8569, 52.7865, 58.1408	342.2787, 47.7526, 76.4114, 97.3304, 14.2178, 163.9513, 94.2575, 280.0273, 16.0971	0.0271
0.2	120.6613, 107.6712, 84.0427, 133.5377, 82.1999, 223.122, 41.1617, 41.9398, 12.016	347.2051, 79.4503, 109.8796, 83.5017, 345.4284, 133.8355, 66.4118, 201.0041, 4.4419	0.0261
0.4*	92.4395, 133.2136, 38.9704, 34.9182, 56.0256, 57.4097, 31.5916, 140.1088, 128.7125	62.036, 16.411, 319.5634, 128.4443, 49.7357, 328.5103, 210.6498, 44.0123, 327.4517	0.0108

tbl:dpsyche_sens

Table 2.2: Results of dPSYCHE optimisations for different sensitivities. Note that since the cost function depends on the value of c, the exact value of $f_{\text{diff},2}$ for these optimisations cannot be compared to one another; they are only presented here for completeness. The final entry, labelled with an asterisk, was run with no limit on the number of function evaluations (this is discussed further in the text).

provided good decoupling in the 'easier' deshielded region and somewhat acceptable performance in the shielded region, with the notable exception of the strongly coupled peaks at 2.5 ppm.

I therefore ran a longer optimisation using the same starting point and with no limit on the number of function evaluations, which successfully lowered the cost function value by twofold (the last entry in table 2.2). The pure shift spectrum obtained using these optimised pulses is shown in fig. 2.21. The decoupling quality in both regions is comparable to that in the PSYCHE spectrum, again with the exception of the strongly coupled peaks at 2.5 ppm.

To make the dPSYCHE experiment more robust towards strong coupling, a TSE version of the sequence was also written and tested (fig. 2.22. This did indeed improve the decoupling at 2.5 ppm, as expected. However, although the TSE-dPSYCHE version still retains its *overall* sensitivity advantage over TSE-PSYCHE (especially evident in the deshielded region), this is not true of every peak: the two peaks at 1.65 and 1.70 ppm seem to have decreased intensities in the TSE-dPSYCHE experiment. The reason for this remains unclear.
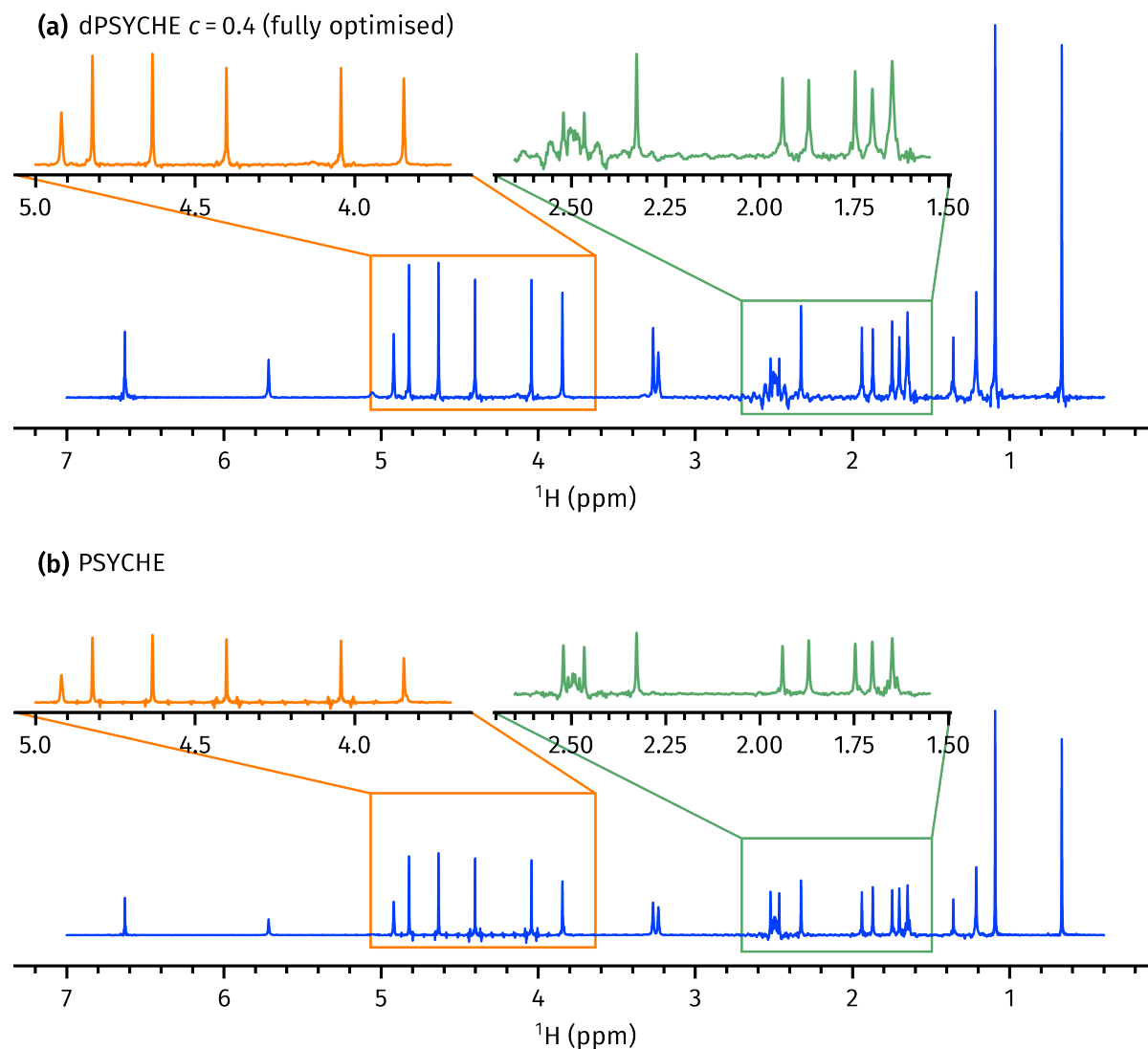
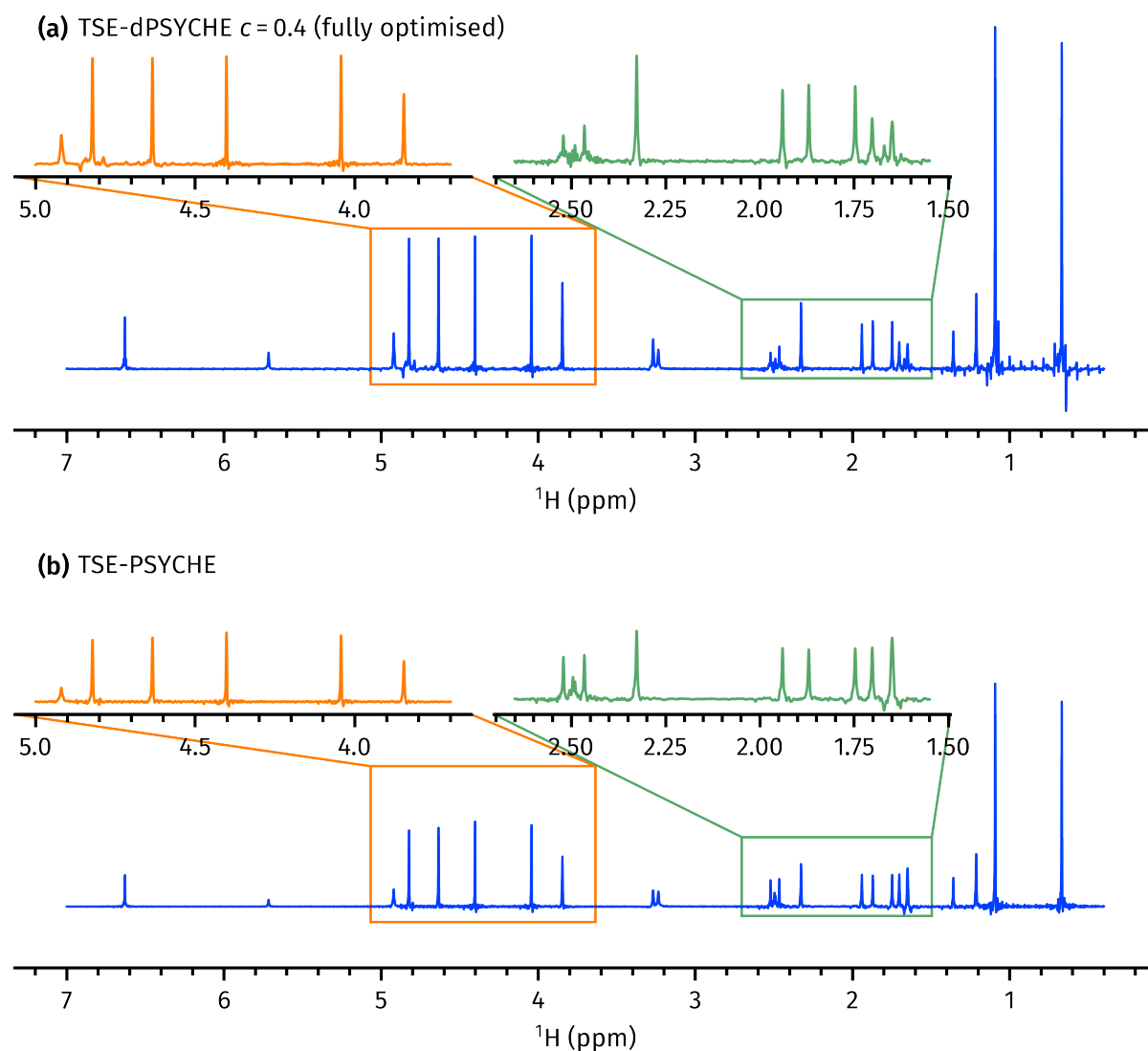*Figure 2.21:* **(a)** Fully optimised dPSYCHE experiment with $c = 0.4$. **(b)** Double-saltire PSYCHE with $\beta = 20°$. *Data code:* `6A-211231`.

2.7   Ultrafast PSYCHE-iDOSY

Collaborative stuff with Jean-Nicolas Dumez which never quite went anywhere (sadly)—although at EUROMAR, JND told me he now had a student (or a postdoc?) working on it and that they might send us some stuff.

2.8   References

(1)   Zangger, K. Pure shift NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2015**, *86-87*, 1–20, DOI: `10.1016/j.pnmrs.2015.02.002`.

Figure 2.22: **(a)** Fully optimised TSE-dPSYCHE experiment with $c = 0.4$. **(b)** Double-saltire TSE-PSYCHE with $\beta = 20°$. *Data code:* `7A-220129`.

Castanar2017MRC   (2)   Castañar, L. Pure shift ^{1}H NMR: what is next? *Magn. Reson. Chem.* **2017**, *55*, 47–53, DOI: `10.1002/mrc.4545`.

Keeler2010   (3)   Keeler, J., *Understanding NMR Spectroscopy*, 2nd ed.; Wiley: Chichester, U.K., 2010.

Thrippleton2005JMR   (4)   Thrippleton, M. J.; Edden, R. A. E.; Keeler, J. Suppression of strong coupling artefacts in J-spectra. *J. Magn. Reson.* **2005**, *174*, 97–109, DOI: `10.1016/j.jmr.2005.01.012`.

Griesinger1986JCP   (5)   Griesinger, C.; Sørensen, O. W.; Ernst, R. R. Correlation of connected transitions by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **1986**, *85*, 6837–6852, DOI: `10.1063/1.451421`.

Aue1976JCP   (6) Aue, W. P.; Karhan, J.; Ernst, R. R. Homonuclear broad band decoupling and two-dimensional J-resolved NMR spectroscopy. *J. Chem. Phys.* **1976**, *64*, 4226–4227, DOI: 10.1063/1.431994.

Xu1991JMR   (7) Xu, P.; Wu, X.-L.; Freeman, R. Broadband-decoupled proton spectroscopy. *J. Magn. Reson.* **1991**, *95*, 132–148, DOI: 10.1016/0022-2364(91)90329-r.

Nuzillard1996JMRSA   (8) Nuzillard, J.-M. Time-Reversal of NMR Signals by Linear Prediction. Application to Phase-Sensitive Homonuclear J-Resolved Spectroscopy. *J. Magn. Reson., Ser. A* **1996**, *118*, 132–135, DOI: 10.1006/jmra.1996.0020.

Simova1997JMR   (9) Simova, S.; Sengstschmid, H.; Freeman, R. Proton Chemical-Shift Spectra. *J. Magn. Reson.* **1997**, *124*, 104–121, DOI: 10.1006/jmre.1996.1001.

Pell2007JMR   (10) Pell, A. J.; Keeler, J. Two-dimensional J-spectra with absorption-mode lineshapes. *J. Magn. Reson.* **2007**, *189*, 293–299, DOI: 10.1016/j.jmr.2007.09.002.

Foroozandeh2015CC   (11) Foroozandeh, M.; Adams, R. W.; Kiraly, P.; Nilsson, M.; Morris, G. A. Measuring couplings in crowded NMR spectra: pure shift NMR with multiplet analysis. *Chem. Commun.* **2015**, *51*, 15410–15413, DOI: 10.1039/c5cc06293d.

Griesinger1985JACS   (12) Griesinger, C.; Soerensen, O. W.; Ernst, R. R. Two-dimensional correlation of connected NMR transitions. *J. Am. Chem. Soc.* **1985**, *107*, 6394–6396, DOI: 10.1021/ja00308a042.

Sorensen1985JACS   (13) Sørensen, O. W.; Griesinger, C.; Ernst, R. R. Time reversal of the evolution under scalar spin-spin interactions in NMR. Application for ω_1 decoupling in two-dimensional NOE spectroscopy. *J. Am. Chem. Soc.* **1985**, *107*, 7778–7779, DOI: 10.1021/ja00311a101.

Oschkinat1986JMR   (14) Oschkinat, H.; Pastore, A.; Pfändler, P.; Bodenhausen, G. Two-dimensional correlation of directly and remotely connected transitions by z-filtered COSY. *J. Magn. Reson.* **1986**, *69*, 559–566, DOI: 10.1016/0022-2364(86)90176-9.

Pell2007MRC   (15) Pell, A. J.; Edden, R. A. E.; Keeler, J. Broadband proton-decoupled proton spectra. *Magn. Reson. Chem.* **2007**, *45*, 296–316, DOI: 10.1002/mrc.1966.

Moutzouri2020ACIE   (16) Moutzouri, P.; Paruzzo, F. M.; Simões de Almeida, B.; Stevanato, G.; Emsley, L. Homonuclear Decoupling in ^{1}H NMR of Solids by Remote Correlation. *Angew. Chem., Int. Ed.* **2020**, *59*, 6235–6238, DOI: 10.1002/anie.201916335.

Bruschweiler2004JCP   (17) Brüschweiler, R.; Zhang, F. Covariance nuclear magnetic resonance spectroscopy. *J. Chem. Phys.* **2004**, *120*, 5253–5260, DOI: 10.1063/1.1647054.

Zhang2004JACS   (18) Zhang, F.; Brüschweiler, R. Indirect Covariance NMR Spectroscopy. *J. Am. Chem. Soc.* **2004**, *126*, 13180–13181, DOI: 10.1021/ja047241h.

Jaeger2014ARNMRS   (19) Jaeger, M.; Aspers, R. L. E. G. Covariance NMR and Small Molecule Applications. *Annu. Rep. NMR Spectrosc.* **2014**, *83*, 271–349, DOI: 10.1016/B978-0-12-800183-7.00005-8.

Morris2010JACS   (20) Morris, G. A.; Aguilar, J. A.; Evans, R.; Haiber, S.; Nilsson, M. True Chemical Shift Correlation Maps: A TOCSY Experiment with Pure Shifts in Both Dimensions. *J. Am. Chem. Soc.* **2010**, *132*, 12770–12772, DOI: 10.1021/ja1039715.

Aguilar2012ACIE (21)   Aguilar, J. A.; Colbourne, A. A.; Cassani, J.; Nilsson, M.; Morris, G. A. Decoupling Two-Dimensional NMR Spectroscopy in Both Dimensions: Pure Shift NOESY and COSY. *Angew. Chem., Int. Ed.* **2012**, *51*, 6460–6463, DOI: 10.1002/anie.201108888.

Foroozandeh2014JACS (22)   Foroozandeh, M.; Adams, R. W.; Nilsson, M.; Morris, G. A. Ultrahigh-Resolution Total Correlation NMR Spectroscopy. *J. Am. Chem. Soc.* **2014**, *136*, 11867–11869, DOI: 10.1021/ja507201t.

Lupulescu2012JMR (23)   Lupulescu, A.; Olsen, G. L.; Frydman, L. Toward single-shot pure-shift solution ^{1}H NMR by trains of BIRD-based homonuclear decoupling. *J. Magn. Reson.* **2012**, *218*, 141–146, DOI: 10.1016/j.jmr.2012.02.018.

Meyer2013ACIE (24)   Meyer, N. H.; Zangger, K. Simplifying Proton NMR Spectra by Instant Homonuclear Broadband Decoupling. *Angew. Chem., Int. Ed.* **2013**, *52*, 7143–7146, DOI: 10.1002/anie.201300129.

Kiraly2018MRC (25)   Kiraly, P.; Nilsson, M.; Morris, G. A. Practical aspects of real-time pure shift HSQC experiments. *Magn. Reson. Chem.* **2018**, *56*, 993–1005, DOI: 10.1002/mrc.4704.

Donovan2014ACIE (26)   Donovan, K. J.; Frydman, L. HyperBIRD: A Sensitivity-Enhanced Approach to Collecting Homonuclear-Decoupled Proton NMR Spectra. *Angew. Chem., Int. Ed.* **2014**, n/a–n/a, DOI: 10.1002/anie.201407869.

Taylor2021MRC (27)   Taylor, D. A.; Natrajan, L. S.; Nilsson, M.; Adams, R. W. SABRE-enhanced real-time pure shift NMR spectroscopy. *Magn. Reson. Chem.* **2021**, *59*, 1244–1252, DOI: 10.1002/mrc.5206.

Aguilar2010ACIE (28)   Aguilar, J. A.; Faulkner, S.; Nilsson, M.; Morris, G. A. Pure Shift ^{1}H NMR: A Resolution of the Resolution Problem? *Angew. Chem., Int. Ed.* **2010**, *49*, 3901–3903, DOI: 10.1002/anie.201001107.

Moutzouri2017CC (29)   Moutzouri, P.; Chen, Y.; Foroozandeh, M.; Kiraly, P.; Phillips, A. R.; Coombes, S. R.; Nilsson, M.; Morris, G. A. Ultraclean pure shift NMR. *Chem. Commun.* **2017**, *53*, 10188–10191, DOI: 10.1039/c7cc04423b.

Zangger1997JMR (30)   Zangger, K.; Sterk, H. Homonuclear Broadband-Decoupled NMR Spectra. *J. Magn. Reson.* **1997**, *124*, 486–489, DOI: 10.1006/jmre.1996.1063.

Garbow1982CPL (31)   Garbow, J. R.; Weitekamp, D. P.; Pines, A. Bilinear rotation decoupling of homonuclear scalar interactions. *Chem. Phys. Lett.* **1982**, *93*, 504–509, DOI: 10.1016/0009-2614(82)83229-6.

Aguilar2011ACIE (32)   Aguilar, J. A.; Nilsson, M.; Morris, G. A. Simple Proton Spectra from Complex Spin Systems: Pure Shift NMR Spectroscopy Using BIRD. *Angew. Chem., Int. Ed.* **2011**, *50*, 9716–9717, DOI: 10.1002/anie.201103789.

Foroozandeh2014ACIE (33)   Foroozandeh, M.; Adams, R. W.; Meharry, N. J.; Jeannerat, D.; Nilsson, M.; Morris, G. A. Ultrahigh-Resolution NMR Spectroscopy. *Angew. Chem., Int. Ed.* **2014**, *53*, 6990–6992, DOI: 10.1002/anie.201404111.

Foroozandeh2018CEJ (34) Foroozandeh, M.; Morris, G. A.; Nilsson, M. PSYCHE Pure Shift NMR Spectroscopy. *Chem. Eur. J.* **2018**, *24*, 13988–14000, DOI: `10.1002/chem.201800524`.

Foroozandeh2020JMR (35) Foroozandeh, M. Spin dynamics during chirped pulses: applications to homonuclear decoupling and broadband excitation. *J. Magn. Reson.* **2020**, *318*, 106768, DOI: `10.1016/j.jmr.2020.106768`.

Kupce1995JMRSB (36) Kupce, E.; Boyd, J.; Campbell, I. D. Short Selective Pulses for Biochemical Applications. *J. Magn. Reson., Ser. B* **1995**, *106*, 300–303, DOI: `10.1006/jmrb.1995.1049`.

Giraud2010ACIE (37) Giraud, N.; Béguin, L.; Courtieu, J.; Merlet, D. Nuclear Magnetic Resonance Using a Spatial Frequency Encoding: Application to J-Edited Spectroscopy along the Sample. *Angew. Chem., Int. Ed.* **2010**, *49*, 3481–3484, DOI: `10.1002/anie.200907103`.

Gubensak2014CC (38) Gubensäk, N.; Fabian, W. M. F.; Zangger, K. Disentangling scalar coupling patterns by real-time SERF NMR. *Chem. Commun.* **2014**, *50*, 12254–12257, DOI: `10.1039/c4cc05892e`.

Mishra2017JMR (39) Mishra, S. K.; Suryaprakash, N. Pure shift edited ultra high resolution NMR spectrum with complete eradication of axial peaks and unwanted couplings. *J. Magn. Reson.* **2017**, *279*, 74–80, DOI: `10.1016/j.jmr.2017.04.014`.

Buchberger2018MRC (40) Buchberger, K.; Walenta, M.; Zangger, K. Extracting unresolved coupling constants from complex multiplets by a real-time J-upscaled SERF experiment. *Magn. Reson. Chem.* **2018**, *56*, 934–940, DOI: `10.1002/mrc.4699`.

Uhrin1993JMRSA (41) Uhrin, D.; Liptaj, T.; Kover, K. E. Modified BIRD Pulses and Design of Heteronuclear Pulse Sequences. *J. Magn. Reson., Ser. A* **1993**, *101*, 41–46, DOI: `10.1006/jmra.1993.1005`.

Sakhaii2009JMR (42) Sakhaii, P.; Haase, B.; Bermel, W. Experimental access to HSQC spectra decoupled in all frequency dimensions. *J. Magn. Reson.* **2009**, *199*, 192–198, DOI: `10.1016/j.jmr.2009.04.016`.

Paudel2013ACIE (43) Paudel, L.; Adams, R. W.; Király, P.; Aguilar, J. A.; Foroozandeh, M.; Cliff, M. J.; Nilsson, M.; Sándor, P.; Waltho, J. P.; Morris, G. A. Simultaneously Enhancing Spectral Resolution and Sensitivity in Heteronuclear Correlation NMR Spectroscopy. *Angew. Chem., Int. Ed.* **2013**, *52*, 11616–11619, DOI: `10.1002/anie.201305709`.

Reinsperger2014JMR (44) Reinsperger, T.; Luy, B. Homonuclear BIRD-decoupled spectra for measuring one-bond couplings with highest resolution: CLIP/CLAP-RESET and constant-time-CLIP/CLAP-RESET. *J. Magn. Reson.* **2014**, *239*, 110–120, DOI: `10.1016/j.jmr.2013.11.015`.

Nolis2019JMR_psHSQC (45) Nolis, P.; Motiram-Corral, K.; Pérez-Trujillo, M.; Parella, T. Broadband homodecoupled time-shared ^{1}H–^{13}C and ^{1}H–^{15}N HSQC experiments. *J. Magn. Reson.* **2019**, *298*, 23–30, DOI: `10.1016/j.jmr.2018.11.005`.

Singh2020JMR (46) Singh, U.; Bhattacharya, S.; Baishya, B. Pure shift HMQC: Resolution and sensitivity enhancement by bilinear rotation decoupling in the indirect and direct dimensions. *J. Magn. Reson.* **2020**, *311*, 106684, DOI: `10.1016/j.jmr.2020.106684`.

Zwahlen1997JACS   (47)   Zwahlen, C.; Legault, P.; Vincent, S. J. F.; Greenblatt, J.; Konrat, R.; Kay, L. E. Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage λ N-Peptide/boxBRNA Complex. *J. Am. Chem. Soc.* **1997**, *119*, 6711–6721, DOI: `10.1021/ja970224q`.

Kupce2007JMR   (48)   Kupče, E.; Freeman, R. Compensated adiabatic inversion pulses: Broadband INEPT and HSQC. *J. Magn. Reson.* **2007**, *187*, 258–265, DOI: `10.1016/j.jmr.2007.05.009`.

Thrippleton2003ACIE   (49)   Thrippleton, M. J.; Keeler, J. Elimination of Zero-Quantum Interference in Two-Dimensional NMR Spectra. *Angew. Chem., Int. Ed.* **2003**, *42*, 3938–3941, DOI: `10.1002/anie.200351947`.

Timari2015CEJ   (50)   Timári, I.; Szilágyi, L.; Kövér, K. E. PSYCHE CPMG–HSQMBC: An NMR Spectroscopic Method for Precise and Simple Measurement of Long-Range Heteronuclear Coupling Constants. *Chem. Eur. J.* **2015**, *21*, 13939–13942, DOI: `10.1002/chem.201502641`.

Koos2016ACIE   (51)   Koos, M. R. M.; Kummerlöwe, G.; Kaltschnee, L.; Thiele, C. M.; Luy, B. CLIP-COSY: A Clean In-Phase Experiment for the Rapid Acquisition of COSY-type Correlations. *Angew. Chem., Int. Ed.* **2016**, *55*, 7655–7659, DOI: `10.1002/anie.201510938`.

Sinnaeve2016ACIE   (52)   Sinnaeve, D.; Foroozandeh, M.; Nilsson, M.; Morris, G. A. A General Method for Extracting Individual Coupling Constants from Crowded ^{1}H NMR Spectra. *Angew. Chem., Int. Ed.* **2016**, *55*, 1090–1093, DOI: `10.1002/anie.201508691`.

Aguilar2018MRC   (53)   Aguilar, J. A.; Belda, R.; Gaunt, B. R.; Kenwright, A. M.; Kuprov, I. Separating the coherence transfer from chemical shift evolution in high-resolution pure shift COSY NMR. *Magn. Reson. Chem.* **2018**, *56*, 969–975, DOI: `10.1002/mrc.4727`.

Kaltschnee2016JMR   (54)   Kaltschnee, L.; Knoll, K.; Schmidts, V.; Adams, R. W.; Nilsson, M.; Morris, G. A.; Thiele, C. M. Extraction of distance restraints from pure shift NOE experiments. *J. Magn. Reson.* **2016**, *271*, 99–109, DOI: `10.1016/j.jmr.2016.08.007`.

Ilgen2021JMR   (55)   Ilgen, J.; Nowag, J.; Kaltschnee, L.; Schmidts, V.; Thiele, C. M. Gradient selected pure shift EASY-ROESY techniques facilitate the quantitative measurement of ^{1}H,^{1}H-distance restraints in congested spectral regions. *J. Magn. Reson.* **2021**, *324*, 106900, DOI: `10.1016/j.jmr.2020.106900`.

Foroozandeh2016ACIE   (56)   Foroozandeh, M.; Castañar, L.; Martins, L. G.; Sinnaeve, D.; Poggetto, G. D.; Tormena, C. F.; Adams, R. W.; Morris, G. A.; Nilsson, M. Ultrahigh-Resolution Diffusion-Ordered Spectroscopy. *Angew. Chem. Int. Ed.* **2016**, *55*, 15579–15582, DOI: `10.1002/anie.201609676`.

Frydman2002PNASUSA   (57)   Frydman, L.; Scherf, T.; Lupulescu, A. The acquisition of multidimensional NMR spectra within a single scan. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15858–15862, DOI: `10.1073/pnas.252644399`.

Pelupessy2003JACS   (58)   Pelupessy, P. Adiabatic Single Scan Two-Dimensional NMR Spectrocopy. *J. Am. Chem. Soc.* **2003**, *125*, 12345–12350, DOI: `10.1021/ja034958g`.

Frydman2003JACS   (59)   Frydman, L.; Lupulescu, A.; Scherf, T. Principles and Features of Single-Scan Two-Dimensional NMR Spectroscopy. *J. Am. Chem. Soc.* **2003**, *125*, 9204–9217, DOI: 10.1021/ja030055b.

Kiraly2021ACIE   (60)   Kiraly, P.; Kern, N.; Plesniak, M. P.; Nilsson, M.; Procter, D. J.; Morris, G. A.; Adams, R. W. Single-Scan Selective Excitation of Individual NMR Signals in Overlapping Multiplets. *Angew. Chem., Int. Ed.* **2021**, *60*, 666–669, DOI: 10.1002/anie.202011642.

Dumez2018PNMRS   (61)   Dumez, J.-N. Spatial encoding and spatial selection methods in high-resolution NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **2018**, *109*, 101–134, DOI: 10.1016/j.pnmrs.2018.08.001.

Nelder1965TCJ   (62)   Nelder, J. A.; Mead, R. A Simplex Method for Function Minimization. *The Computer Journal* **1965**, *7*, 308–313, DOI: 10.1093/comjnl/7.4.308.

McKinnon1998SIAMJO   (63)   McKinnon, K. I. M. Convergence of the Nelder–Mead Simplex Method to a Nonstationary Point. *SIAM J. Optim.* **1998**, *9*, 148–158, DOI: 10.1137/s1052623496303482.

Hogben2011JMR   (64)   Hogben, H. J.; Krzystyniak, M.; Charnock, G. T. P.; Hore, P. J.; Kuprov, I. Spinach – A software library for simulation of spin dynamics in large spin systems. *J. Magn. Reson.* **2011**, *208*, 179–194, DOI: 10.1016/j.jmr.2010.11.008.

Khaneja2005JMR   (65)   Khaneja, N.; Reiss, T.; Kehlet, C.; Schulte-Herbrüggen, T.; Glaser, S. J. Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms. *J. Magn. Reson.* **2005**, *172*, 296–305, DOI: 10.1016/j.jmr.2004.11.004.

deFouquieres2011JMR   (66)   De Fouquieres, P.; Schirmer, S. G.; Glaser, S. J.; Kuprov, I. Second order gradient ascent pulse engineering. *J. Magn. Reson.* **2011**, *212*, 412–417, DOI: 10.1016/j.jmr.2011.07.023.

Glaser2015EPJD   (67)   Glaser, S. J.; Boscain, U.; Calarco, T.; Koch, C. P.; Köckenberger, W.; Kosloff, R.; Kuprov, I.; Luy, B.; Schirmer, S.; Schulte-Herbrüggen, T.; Sugny, D.; Wilhelm, F. K. Training Schrödinger's cat: quantum optimal control. *Eur. Phys. J. D* **2015**, *69*, No. 279, DOI: 10.1140/epjd/e2015-60464-1.

Goodwin2016JCP   (68)   Goodwin, D. L.; Kuprov, I. Modified Newton-Raphson GRAPE methods for optimal control of spin systems. *J. Chem. Phys.* **2016**, *144*, 204107, DOI: 10.1063/1.4949534.

Geen1991JMR   (69)   Geen, H.; Freeman, R. Band-selective radiofrequency pulses. *J. Magn. Reson.* **1991**, *93*, 93–141, DOI: 10.1016/0022-2364(91)90034-q.

Emsley1990CPL   (70)   Emsley, L.; Bodenhausen, G. Gaussian pulse cascades: New analytical functions for rectangular selective inversion and in-phase excitation in NMR. *Chem. Phys. Lett.* **1990**, *165*, 469–476, DOI: 10.1016/0009-2614(90)87025-m.

Ewing1990CP   (71)   Ewing, B.; Glaser, S. J.; Drobny, G. P. Development and optimization of shaped NMR pulses for the study of coupled spin systems. *Chem. Phys.* **1990**, *147*, 121–129, DOI: 10.1016/0301-0104(90)85028-u.

Higham2005SIAMJMAA (72) Higham, N. J. The Scaling and Squaring Method for the Matrix Exponential Revisited. *SIAM J. Matrix Anal. Appl.* **2005**, *26*, 1179–1193, DOI: 10.1137/04061101x.

Moler2003SIAMR (73) Moler, C.; Van Loan, C. Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later. *SIAM Rev.* **2003**, *45*, 3–49, DOI: 10.1137/s00361445024180.

Kelley1999 (74) Kelley, C. T., *Iterative Methods for Optimization*; Society for Industrial and Applied Mathematics: 1999, DOI: 10.1137/1.9781611970920.

Nocedal2006 (75) Nocedal, J.; Wright, S. J., *Numerical Optimization*, 2nd ed.; Springer: New York, 2006.

refsection:3

refsection:4

List of figures

List of tables

refsection:5