Optimization of Magnetization Transfer Experiments for Kinetic Rate Measurements

Haim Taitelbaum,†‡ George H. Weiss‡* and Richard G. S. Spencer§
†Division of Computer Research and Technology, National Institutes of Health, Bethesda, MD 20892, USA
‡Jerome Schottenstein Cellscan Center and Department of Physics, Bar-Ilan University, Ramat-Gan 52100, Israel
§Laboratory of Cellular and Molecular Biology, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA

A framework is developed for optimizing magnetization transfer experiments for measuring first-order rate constants of a two-component exchange network. In the analysis the spin-lattice relaxation time of the non-saturated resonance is assumed to be known. However, we demonstrate that our work can also be applied to cases in which the spin-lattice relaxation time falls within a physiologically reasonable range for this parameter. We provide an algorithm for the parameters permitting the experimenter to optimize the selection of measurement times in order to control the error in a worst-case scenario. In addition, we compare the performance of the time-dependent saturation experiment with an inversion experiment in this framework, finding that the former always leads to a more effective design.

INTRODUCTION

There are a number of optimization procedures for selecting the time points at which data are taken in measurements of spin-lattice relaxation times \( T_1 \) in chemical systems. Some of these are summarized in a review by Weiss and Ferretti. However, similar issues are raised by the measurement of reaction rates in physiological systems. An interesting discussion of this problem based on computer simulations together with a comparison of different techniques available for estimating the reaction rate has been given by Rydzy et al.

In the present paper we adopt a different approach to the problem, which yields more general results. Here we analyze the design of time-dependent saturation-transfer (TDST) sequences and inversion-recovery (IR) experiments by adapting a general formalism for optimal design of \( T_1 \) measurements originally developed in Ref. 18 and reviewed in Ref. 1. The term "optimization" refers to an experimental design that minimizes the upper bound on the error of the estimate of the reaction rate through the choice of the set of time points at which magnetization measurements are made. Since the rate constants are \textit{a priori} unknown, our results cannot be fully optimal for all rate constants. However, our results provide an upper bound on the error due to noise which is independent of the rate constants, provided that they lie in a specified interval.

The system to be considered consists of two species, A and B, which interconvert according to:

\[
A \underset{k'}{\overset{k}{\rightleftharpoons}} B
\]  
(1)

where \( k \) and \( k' \) denote rate constants that characterize the kinetics of the indicated pseudo-first-order reaction. The spin-lattice relaxation times of the two species will be denoted by \( T_1(A) \) and \( T_1(B) \), respectively, which are assumed to be known in our analysis. This assumption is generally validated by information from other experiments which provide approximate values. Alternatively, the system may be perturbed in such a way that the relative reaction rates are altered to a much greater extent than the relative relaxation times, e.g., as for muscle stimulation.

We note that the two-site exchange network modeled by Eqn (1) with a single \( T_1 \) for each species is likely to be an approximation. Nevertheless, this model has been used extensively to describe the creatine kinase catalyzed transfer of high-energy phosphate between phosphocreatine and ATP in muscle and brain.

Extensions of the exchange network to account for more than two sites have been studied in Refs 7 and 8.

We now define the relation between the saturation time \( (TS) \) and the recycle delay \( (RD) \), that is, the time between the observation pulse and the following period of saturation (Fig. 1). In order to maintain the same degree of partial saturation of non-saturated resonances for each choice of saturation time we perform the experiment with variable RDs to compensate for the variable TS. Thus, on denoting by \( TS \) and \( RD \) the saturation time and recycle delay for the \( i \)th measurement time point, respectively, we impose the constraint:

\[
TS_i + AQ_i + RD_i = T_c
\]  
(2)

where \( AQ \) is the constant acquisition time and \( T_c \) is the total cycle time of the pulse sequence which is selected in accordance with the degree of spectral saturation acceptable to the experimenter. The \( TS \) are given by the algorithm whose design we will be discussing, and the \( RD \) are then calculated from Eqn (2). Note that, in contrast to the analyses given in Refs 18–20, the duration of the experiment is independent of the selection of the \( TS \), depending only on \( T_1 \). The total experimental time is fixed and equal to \( mT_1 \), where \( m \) is the number of time points (typically between 6 and 12) and \( n \) is the number of acquisitions at each time point.
ANALYSIS

In these experiments one species is saturated, say the B species, and the magnetization of the A species, $M_A(t)$, is measured as a function of time. For a noise-free system this function is the solution to:

$$\frac{dM_A(t)}{dt} = \frac{M_0 - M_A(t)}{T_1(A)} - kM_A(t)$$  \hspace{1cm} (3)

Different classes of experiments are distinguished by the choice of initial conditions. We will be interested primarily in the TDST experiment for which $M_A(0) = M_0$. If $\epsilon(t)$ denotes instrumental noise we have:

$$M_A(t) = \frac{M_0}{1 + kT_1(A)}[1 + kT_1(A)e^{-\tau t}] + \epsilon(t)$$  \hspace{1cm} (4)

where $\tau$, the apparent spin-lattice relaxation time, is given by:

$$\frac{1}{\tau} = \frac{1}{T_1(A)} + k$$  \hspace{1cm} (5)

The noise will be assumed to have a zero mean and to be uncorrelated at the different measurement times:

$$\langle \epsilon(t_i) \rangle = 0 \quad \langle \epsilon(t_i)\epsilon(t_j) \rangle = \sigma_{\text{noise}}^2 \delta_{i,j}$$  \hspace{1cm} (6)

where the angled brackets refer to an average with respect to the distribution of the noise, $t_i$ is the $i$th time point, and $\delta_{i,j}=0$ when $i=j$ and $=1$ when $i=j$. The ratio $M_0/\sigma_{\text{noise}}$ will be used as a measure of the signal to noise ratio (SNR).

Although only the rate constant $k$ is of direct interest it is also necessary to estimate the constant $M_0$ from the experimental data. The estimates of these quantities will be denoted by $\hat{k}$ and $\hat{M}_0$, respectively. They are random variables that depend on experimental noise. In our analysis we consider only the case in which $M_0$ and $k$ are estimated from the same set of measurements. It is also possible to analyze experiments in which $M_0$ is estimated separately.\(^{18}\) However, this generally requires longer running times to achieve a target precision or, for a fixed running time, leads to a less precise estimate of the rate constant $k$.

We require initial information related to the range of the parameter to be estimated in analogy to the analysis reviewed in Ref. 1. This will be expressed in the form:

$$k_{\text{min}} \leq k \leq k_{\text{max}}$$  \hspace{1cm} (7)

It has been shown in Ref. 18, that specifying such an initial uncertainty interval rather than an initial estimate of the value of the rate constant results in much less drastic effects on the precision of the estimate when the assumption is violated. For convenience we define the dimensionless parameter:

$$\beta = \frac{k_{\text{max}}}{k_{\text{min}}}$$  \hspace{1cm} (8)

Typical values of $\beta$ for in vivo experiments might range from 5 to 10. More sophisticated experiments can be designed in which the uncertainty interval is changed adaptively as more data are acquired,\(^{19,20}\) but in the interest of simplicity we will not consider this additional generalization here.

In the present investigation we consider only the case in which measurements of $M_A(t)$ are made at a set of $n$ time points, $\{t_i\}, i=1,2, \ldots, n$, and the parameters are estimated using a linearized least squares analysis. Let the functional form of the magnetization at time $t_i$ in the absence of noise be denoted by $h_i(M_0,k)$ so that the measured magnetization at time $t_i$ may be written $M_A(t_i) = h_i(M_0,k) + \epsilon(t_i)$. If $n$ is fixed we seek to minimize the function:

$$\varphi(\hat{M}_0, \hat{k}) = \sum_{i=1}^{n} [M_A(t_i) - h_i(\hat{M}_0, \hat{k})]^2$$  \hspace{1cm} (9)

by an appropriate choice of the set of measurement times $\{t_i\}$. We will assume that the SNR is sufficiently large so that in an expansion of the difference $h_i(M_0,k) - h_i(M_0, k)$ only first-order terms in $\delta M_0 = M_0 - \hat{M}_0$ and $\delta k = k - \hat{k}$ need be retained. It has been shown,\(^{18}\) in the context of measurements of $T_1$, that this is a reasonable approximation for SNR$\geq7$. Let $\langle (\delta k)^j \rangle$ denote the $j$th moment of the random variable $\delta k$. It is easy to verify that if only the lowest order terms in $\delta k$ are retained then $\langle \delta k \rangle = 0$, i.e. the estimate of $k$ is unbiased in this approximation. The function to be
minimized in our analysis is the normalized standard deviation of the rate constant \(o(k)/k = (\langle \delta k \rangle^2)^{1/2} / k\) which we denote by \(s_s(k)\). To find an expression for this function we follow the linearized analysis in Ref. 18, obtaining:

\[
s_s(k) = \frac{\sigma_{\text{noise}}}{M_s T_1} \left( \frac{W}{UW-V^2} \right)^{1/2} = \frac{\sigma_{\text{noise}}}{M_0} G_s(k; \{t\}) \tag{10}\]

in which \(U\), \(V\) and \(W\) are dimensionless variables defined by:

\[
U = \left[ \frac{M_0 T_1(A)}{\sigma_{\text{noise}}} \right] \sum_{i=1}^{n} \left[ \frac{\partial M_A(t_i)}{\partial k} \right]^2
\]

\[
V = \frac{1}{M_0 T_1(A)} \sum_{i=1}^{n} \left[ \frac{\partial M_A(t_i)}{\partial M_0} \right]
\]

\[
W = \sum_{i=1}^{n} \left[ \frac{\partial M_A(t_i)}{\partial M_0} \right]^2 \tag{11}\]

Since the coefficient of \(G_s(k;\{t\})\) is equal to the reciprocal of the \(S/N\), which is independently measurable, we need only minimize the function \(G_s(k;\{t\})\) with respect to the \(t_i\).

The parameter \(k\) is an argument in \(G_s(k;\{t\})\), so that the normalized standard deviation presumably depends on the parameter to be estimated. Since we cannot know this parameter initially, we adopt an alternative strategy, again following Ref. 18, of choosing the set of \(t_i\) that will minimize the function:

\[
\max_{(t_{\text{max}}, k_{\text{max}})} G_s(k; \{t\})
\]

where the indicated maximization is to be carried out over all \(k\) in the interval \((k_{\text{min}}, k_{\text{max}})\). This minimum will be denoted by \(G_s^* (k; \{t\})\):

\[
G_s^* (k; \{t\}) = \min_{\{t\}} \left[ \max_{(k_{\text{min}}, k_{\text{max}})} G_s(k; \{t\}) \right] \tag{12}\]

which is a firm upper bound on the error.

For simplicity we minimize \(G_s(k;\{t\})\) only over sets of \(t_i\) generated according to a specific rule. We consider a class of the \(t_i\) defined by three parameters, \(t_{\text{min}}, t_{\text{max}}\) and \(r\), where \(t_{\text{max}}\) is the time of the earliest measurement, \(t_{\text{min}}\) is the latest measurement time and \(r\) defines a polynomial through the relation:

\[
t_i = t_{\text{min}} + (t_{\text{max}} - t_{\text{min}}) \left( \frac{i-1}{n-1} \right)^r \quad (i = 1, 2, \ldots, n) \tag{13}\]

In Fig. 2 we show the effect of varying the parameter \(r\) in Eqn (13), on the value of \(G_s^* (k;\{t\})\) regarded as a function of \(\beta\). The results suggest that a linear spacing of time points is generally not a good design choice according to our criterion. We interpret these results as indicating an advantage to concentrating the \(t_i\) at early times when the SNR is greater and \(M_s(t)\) changes most rapidly, thereby providing greater sensitivity to the variation of parameters.

In calculating the optimal spacing we also need to consider the variation of \(t_s = t_{\text{max}}\) as a function of \(\beta\) since decreasing this parameter shortens the experiment. Figure 3 shows a typical plot of the dimensionless time \(r_s = t_s / T_1(A)\) as a function of \(\beta\) for \(n = 12\) and \(r = 1.2\) and 3. We see that \(r = 2\) leads to a smaller value of \(t_{\text{max}}\) than does \(r = 3\). Quite generally quadratic spacing seems to be superior to linear spacing, in terms of \(G_s^* (k;\{t\})\), and comparable to cubic spacing, but it also has a shorter total running time than a design with \(r = 3\). In Table 1 we therefore summarize the optimal values of \(t_{\text{max}}\) for \(r = 2\). Notice that when the number of time points is relatively small for small \(\beta\), the value of \(t_s\) can be rather large. This is because the small initial uncertainty allows measurements to be made at times which otherwise might furnish little information. As \(n\) increases long-time measurements are not required to estimate \(k\). When \(n\) is held fixed and \(\beta\) allowed to increase, the value of \(t_s\) decreases. If, on the other hand, \(\beta\) is held fixed and \(n\) allowed to increase there is a crossover between decreasing and increasing values of \(t_s\).

Figure 4 indicates the behavior of \(G_s^* (k;\{t\})\) considered as a function of the dimensionless rate constant \(k = k T_1(A)\). Optimal sets of the \(t_i\) have been used to generate the curves. It should be borne in mind that \(k\) does not necessarily lie in the interval \((k_{\text{min}}, k_{\text{max}})\) since the interval only represents an educated guess. Hence the function \(G_s^* (k;\{t\})\) should be evaluated on a larger interval than is implied by \(\beta\). The major point that emerges from Fig. 4 is that there is a significant advantage to choosing a quadratic or cubic as opposed to a linear spacing of measurement times.

We have also considered the IR experiment, which is performed by broad-band inversion followed by the saturation of one resonance for a time \(T_s\), an observe pulse, data acquisition, and finally a recycle delay \(R_D\) as in Fig. 1. A potential advantage of the IR experiment is its increased dynamic range which is advantageous in estimating the parameter \(M_0\). While this is not of direct interest, it plays a role in determining the

\[
\text{Figure 2. Curves of } G_s^* \text{ plotted as a function of } \beta \text{ for } n = 12 \text{ and } r = 1.2 \text{ and 3.}
\]
precision of the estimate of $k$ because the estimates of these two parameters are correlated. To analyze the IR experiment Eqn (3) is to be solved subject to the initial condition $M_A(0) = -M_0$. Equation (4) is then replaced by:

$$M_A(t) = \frac{M_0}{1 + kT(A)} \left[1 - \left(2 + kT(A)\right)e^{-rt}\right] + \epsilon(t) \quad (14)$$

The least-squares formalism as outlined in Eqns (10) and (11) remains unchanged with the derivatives required for the evaluation of Eqn (11) calculated from Eqn (14). To compare the performance of TDST and IR, we evaluated the bound on the normalized standard deviation when both the number of measurements and the cycle time are held fixed. Our analysis demonstrates that the TDST experiment always gives a bound lower than the comparable IR experiment. Hence we recommend that the TDST experiment be used in preference to the IR design. This conclusion, of course, is based on the specific criterion given in Eqn (12) and must be examined with some care if some other criterion is used.

It is interesting to compare the results produced by the present theory with a design used to measure the rate constant in a study of the kinetics of phosphate exchange in the perfused rat heart. We take, as a typical value suggested by Ref. 8, $T(A) = 3$ s and the a priori range of $k$ to be from 0.5 to 2.5 s$^{-1}$ which implies that $\beta = 5$. In these experiments $n = 6$ measurements were taken at the time points $t_i = 0, 0.3, 0.6, 1.2, 2.4$ and 4.8 s which represents geometric spacing with a ratio of 2. We also include a further comparison with Ref. 17 in which an arithmetic progression of time points is used, that is $t_{i+1} = t_i + (i + 1)\Delta t$, where $\Delta t$ is a constant interval. In Fig. 5 we show the results obtained for the function $G_n$ over a range of values of the dimensionless parameter $\kappa$. The solid line gives the bound derived from the present theory using optimal quadratic spacing, the dashed curve is calculated using the six time points used in Ref. 8 and the dotted curve is calculated from a typical set of the kind used in Ref. 17. The curves in Fig. 5(a) indicate that the parameter spacing in both Refs 8 and 17 lead to a lower standard deviation over much of the range considered. While this finding might appear to favor designs based on geometric spacing it should be emphasized that our class of designs provides an upper bound on the measure of precision for all possible values of the dimensionless rate $\kappa$. This does not rule out the possibility that some other design might be more effective in a specific experiment, as is indeed seen to be the case. In Fig. 5(b) we show a similar comparison with $T(A) = 1$ s which shifts the curves to a different range in $\kappa$. Here, our algorithm based on $r = 2$ yields a lower standard deviation than do the other methods over a large portion of the range.

We comment also that our formalism can also furnish an estimate of the normalized standard deviation derived from measurements made at any set of measurement times, provided that a linearized analysis is reasonable, i.e., provided that the SNR is sufficiently large. In such an analysis, one inserts the estimated rate constant $\dot{k}$ into Eqn (10) to find the standard deviation after the measurements have been made. If, in addition, the noise $\epsilon(t)$ is assumed to have a Gaussian distribution then one can also provide confidence limits on the estimate. It was shown in Ref. 17 that an SNR greater than 7 sufficed to validate the use of a linearized analysis in designing experiments to measure $T_i$. We have not carried out analogous calculations for the

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.95</td>
<td>2.62</td>
<td>2.53</td>
<td>2.49</td>
</tr>
<tr>
<td>3</td>
<td>26.74</td>
<td>2.62</td>
<td>2.53</td>
<td>2.49</td>
</tr>
<tr>
<td>5</td>
<td>2.80</td>
<td>2.62</td>
<td>2.53</td>
<td>2.49</td>
</tr>
<tr>
<td>6</td>
<td>2.80</td>
<td>2.62</td>
<td>2.53</td>
<td>2.49</td>
</tr>
<tr>
<td>7</td>
<td>2.80</td>
<td>2.49</td>
<td>2.53</td>
<td>2.49</td>
</tr>
<tr>
<td>8</td>
<td>2.80</td>
<td>2.07</td>
<td>2.13</td>
<td>2.16</td>
</tr>
<tr>
<td>9</td>
<td>2.80</td>
<td>1.81</td>
<td>1.86</td>
<td>1.88</td>
</tr>
<tr>
<td>10</td>
<td>2.80</td>
<td>1.63</td>
<td>1.67</td>
<td>1.70</td>
</tr>
</tbody>
</table>
While our mathematical analysis has assumed as known the value of $T_1(A)$ the results can also be applied to cases in which it is not known. For such cases a best estimate of $T_1(A)$ is used instead of a known value. An analysis which accounts for the unknown spin-lattice relaxation time is more complicated and will be considered in a later publication. It may also be worthwhile to consider spacing patterns other than the class summarized in Eqn (13), as suggested by the utility of the geometric spacing pattern suggested by the curves in Fig. 5.

Finally, the designs considered here specify that the number of measurements are fixed and the allocation of time points is to be optimized. A related question is that of determining the number of time points, $n$, which should be used for estimating $k$ given some form of time constraint. The work of Box and Lucas\textsuperscript{23} and Atkinson and Hunter\textsuperscript{24} suggests that it is optimal to use only as many distinct times as there are parameters to be estimated so that each point will be estimated as accurately as possible. This means that if the number of parameters is equal to $j$ and $l$ measurements are to be made at each time point, these should be allocated equally with $m = l/j$ measurements made at each of $n = j$ measurement times. While the specifics of our NMR experiments do not quite correspond to the hypotheses required in Refs 20 and 21 we have made further calculations of $G_j^*(k)$ as a function of $n$ which tend to support the conclusion that as few as possible distinct measurement points should be used, while the total number of measurements being allocated equally among them. In the case of present interest there are two parameters to be estimated, $k$ and $M_0$, so that the theory of Refs 21 and 22 as well as our own analysis suggests that a completely optimal experiment should consist of measurements made at two points only. Such a design has a serious shortcoming in not allowing the experimenter to verify that the data are consistent with the model. In such a situation it is left to the experimenter to decide how many measurement times suffice to offer satisfactory evidence that the model is a reasonable one, following which our analysis may be applied.

REFERENCES


