Journal of Magnetic Resonance 195 (2008) 45–51

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Polynomially scaling spin dynamics II: Further state-space compression using Krylov subspace techniques and zero track elimination

Ilya Kuprov *

Department of Chemistry, University of Durham, South Road, Durham DH1 3LE, UK

ARTICLE INFO

Article history:
Received 14 June 2008
Revised 14 August 2008
Available online 27 August 2008

Keywords:
NMR
EPR
Spin
Simulation
Scaling

ABSTRACT

We extend the recently proposed state-space restriction (SSR) technique for quantum spin dynamics simulations [Kuprov et al., J. Magn. Reson. 189 (2007) 241–250] to include on-the-fly detection and elimination of unpopulated dimensions from the system density matrix. Further improvements in spin dynamics simulation speed, frequently by several orders of magnitude, are demonstrated. The proposed zero track elimination (ZTE) procedure is computationally inexpensive, reversible, numerically stable and easy to add to any existing simulation code. We demonstrate that it belongs to the same family of Krylov subspace techniques as the well-known Lanczos basis pruning procedure. The combined SSR + ZTE algorithm is recommended for simulations of NMR, EPR and Spin Chemistry experiments on systems containing between 10 and 10^4 coupled spins.

1. Introduction

The current state of magnetic resonance theory can be described as comfortable—most pulse sequences and spin dynamics experiments in both solid and liquid state can be simulated with high accuracy [1–8], at least numerically [1,3,6–8] and in many cases analytically [9,10]. The only hard limit is the available computing power—while many approximations succeed in reducing computation time by large factors [8,11–13], and many special cases can be dealt with efficiently [14–16], the asymptotic scaling is in most cases exponential [11] and accurate simulations of arbitrarily coupled systems with more than ~10 spins are difficult to perform.

We recently proposed a solution to the exponential scaling problem. Our adaptive state-space restriction (SSR) algorithm [17] analyzes the spin interaction graph and generates a restricted state space, keeping only the states that are likely to contribute to system evolution. Dramatic reduction in simulation times was achieved and polynomial scaling demonstrated, but we did observe at the time that the reduced operators and state vectors were still very sparse, suggesting that further reduction is possible. In this communication we achieve this reduction with the aid of further algorithmic steps, based on Krylov subspace techniques [18,19]. The operator matrices and state vectors resulting from the procedures described below are dense (and tiny), even for systems with hundreds of spins, meaning that the basis pruning is complete and a computationally efficient Minimal Truncation Scheme [13,20] for spin dynamics simulations has finally been found.

2. Trajectory level pruning and Krylov subspaces

The SSR procedure [17] constructs the reduced basis by analyzing the spin interaction topology and performing what may be termed *spin system level pruning*—it selects the states that can potentially contribute to spin system evolution under a variety of pulse sequences. In practical calculations, however, there is also a considerable scope for optimizations that are specific to the experiment being simulated and to the actual trajectory that the spin system describes. Therefore, a deeper *trajectory level pruning* can be performed for each specific simulation instance to further select the states that do actually contribute to the spin system evolution.

The trajectory level pruning idea is illustrated schematically in Fig. 1—even in a restricted state space, the system trajectory at any given period is usually confined to some small subspace (which happens to be Krylov subspace [18,19], as we will see from the algebraic analysis below). On a practical level, this may be inferred from direct inspection of magnetic resonance simulation trajectories and invariant subspaces corresponding to different irreducible representations of system symmetry groups [21–23]. Similarly, product operator treatments of even very sophisticated pulse sequences only ever require small operator basis sets [24]. The Krylov subspaces in Fig. 1 are switched (shown as intersections between planes) or modified by external events, such as pulses or decoupling, and a sequence of these subspaces contains the system

* Corresponding author. Fax: +44 191 334 2216.
E-mail address: ilya.kuprov@durham.ac.uk

1090-7807/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2008.08.008

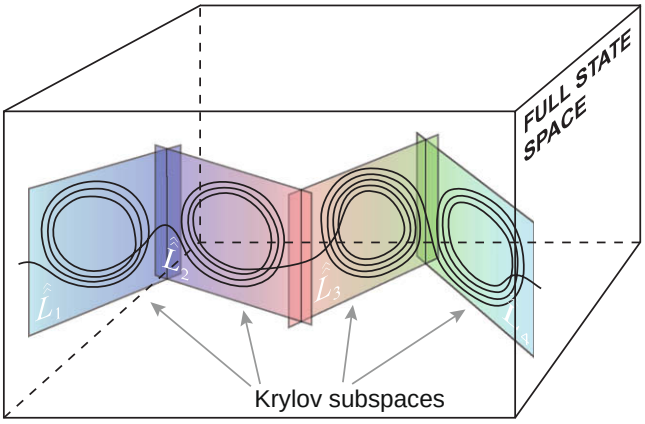


Fig. 1. Schematic illustration of the trajectory level state-space pruning. The analysis of practical simulations as well as theoretical arguments given in the main text suggest that the spin system trajectory is usually confined to relatively small subspaces (sketched above as a 2D planes) of the full state space. Switches between the Liouvillian operators caused by the pulse sequence events move the trajectory between these subspaces, but they do stay small. A computationally efficient algorithm for mapping these subspaces (which are shown below to be Krylov subspaces) would reduce the dimension of the matrices involved in the simulation, leading to an improvement in speed and scaling.

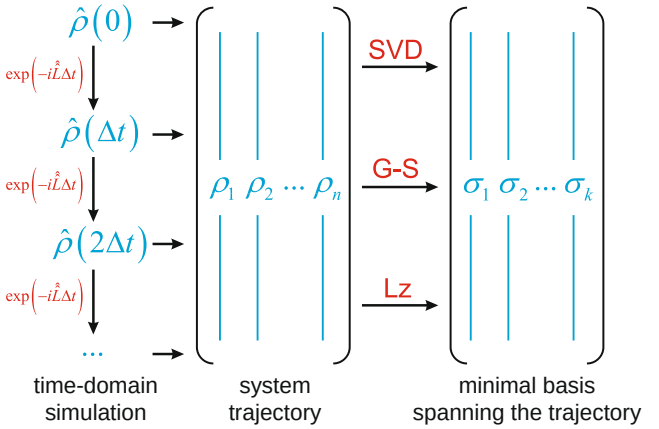


Fig. 2. Schematic flowchart of trajectory level pruning. The state vectors $\{\rho_1, \rho_2, \ldots, \rho_n\}$ comprising the system trajectory under a given Liouvillian $\hat{\hat{L}}$ are, in general, linearly dependent and belong to a subspace of the full state space. The minimal basis $\{\sigma_1, \sigma_2, \ldots, \sigma_k\}$ spanning this trajectory may be generated using one of the standard orthogonalization techniques: Gram-Schmidt (G-S) [28], singular value decomposition (SVD) [29] and Lanczos (Lz) [13,20]. It is known [13,20] that the dimension of this minimal basis is significantly smaller than the full state space dimension.

trajectory. If a computationally efficient way can be found to map out those subspaces, the size of the simulation is likely to be reduced.

While the spin system level pruning and the resulting polynomially scaling simulation algorithm are a recent development [10], the trajectory level pruning (which, on its own, scales exponentially[1]), has a long history, starting with the Lanczos/Arnoldi procedure proposed by Freed et al. [13,20] and the general Krylov subspace techniques used in molecular spectroscopy, electrical circuit analysis and control systems theory [18,19]. We will now outline the connection between the algebraic operations performed in a spin dynamics simulation and the general theory of Krylov subspaces.

In a typical time-domain simulation (and the sophisticated modern pulse sequences do increasingly require time-domain simulations), a system trajectory under a given Liouvillian $\hat{\hat{L}}$ is generated and a scalar product of that trajectory with the observable state is computed to obtain the evolution curve for the observable we are interested in. Formally speaking, the objective is to generate the propagator group orbit $G(\hat{\rho}_0)$ of the initial state vector $\hat{\rho}_0$ [17]

$$G(\hat{\rho}_0) = \{\mathrm{e}^{-i\hat{\hat{L}}t}\hat{\rho}_0,\ t \in [0,\infty)\} \tag{1}$$

This orbit is known in magnetic resonance under the name of the 'system trajectory'. In a discrete simulation with a time step Δt the orbit will be discretized as well:

$$G(\hat{\rho}_0) = \{\hat{\rho}_0, \hat{\hat{P}}\hat{\rho}_0, \hat{\hat{P}}^2\hat{\rho}_0, \ldots\};\ \hat{\hat{P}} = \mathrm{e}^{-i\hat{\hat{L}}\Delta t} \tag{2}$$

The pruning problem consists in finding (exactly or approximately) the minimal basis set that contains the entire trajectory. This problem is known as Krylov subspace construction [18,19] and the corresponding space

$$K_n = \mathrm{span}\{\hat{\rho}_0, \hat{\hat{P}}\hat{\rho}_0, \hat{\hat{P}}^2\hat{\rho}_0, \ldots, \hat{\hat{P}}^{n-1}\hat{\rho}_0\} \tag{3}$$

is known as the n-th Krylov subspace generated by the propagator $\hat{\hat{P}}$ and the initial state vector[2] $\hat{\rho}_0$. Because many spin states are not populated in typical simulations [17], the dimension of the Krylov subspace is frequently much smaller than the dimension of the full state space, and a suitably orthogonalized sequence of state vectors in Eq. (3) provides a much smaller basis set while keeping the simulation essentially exact. This process is schematically illustrated in Fig. 2—the system trajectory $\{\hat{\rho}_n\}$ generated by repeated application of the temporal propagation operator to the initial state vector is orthogonalized using one of the standard techniques and the resulting orthonormal set $\{\hat{\sigma}_n\}$ is used as a new basis set. The minimal basis construction is considered complete when the next trajectory vector is contained entirely (or within a user-specified tolerance on the residual norm) in the Krylov subspace generated at the previous step.

The pruning flowchart as given in Fig. 2 is only practical for the smallest spin systems. With more than 15 spins, computing the matrix exponential is not an option due to astronomical CPU and memory requirements [25,26]. Still, it is useful for the formal analysis presented below. Because the Taylor series for $\hat{\hat{P}} = \mathrm{e}^{-i\hat{\hat{L}}t}$ converges for any finite value of t, the Krylov subspace in Eq. (3) is identical to the following Krylov subspace:

$$K_n = \mathrm{span}\{\hat{\rho}_0, \hat{\hat{L}}\hat{\rho}_0, \hat{\hat{L}}^2\hat{\rho}_0, \ldots, \hat{\hat{L}}^m\hat{\rho}_0\} \tag{4}$$

(where m can be much larger than n) because those powers of the Liouvillian do appear in the Taylor expansion of the exponential. The Krylov subspace in the Eq. (4), however, is much simpler to compute—it only needs matrix-vector multiplication and an orthogonalization pass (possibly with elimination of linear dependence), resulting in the required minimal basis set. This is known as Lanczos pruning (since Lanczos procedure provides a convenient orthogonalization and Liouvillian tridiagonalization technique), which Moro and Freed originally formulated in a purely mathematical iterative fashion [13]. It may be shown to be related to the time-dependent perturbation theory, where different powers of the Liouvillian represent perturbative corrections. An important difference from the implementations existing in computer science is that the starting vector is not generated randomly, but is actually the phys-

[1] All pruning methods based on matrix factorizations and model order reduction share one critical weakness—the exact Liouvillian has to be stored and manipulated in order to be reduced. Liouvillian dimensions scale exponentially with the spin system size and with over 15 spins the reduction procedure inevitably runs out of storage space, even with sparse matrices.

[2] Density matrices will mostly be referred to as 'vectors' henceforth, because of the way such calculations are usually implemented—the density matrix written in a suitable basis is stretched into a vector and multiplied by the propagator matrix.

ically correct starting vector for the spin dynamics simulation in question.

Other trajectory level pruning algorithms differ in the way that the orthogonalization is carried out in Eq. (4)—Gram-Schmidt or singular value decomposition (SVD) techniques may be used instead. The Lanczos type methods are relatively fast because the orthogonalization is performed within the algorithm step by projecting out the subspace obtained in the previous iterations [13,20]

$$\beta_k|k\rangle = \left(1 - \sum_{j=1}^{k-1} |j\rangle\langle j|\right)\hat{\hat{L}}|k-1\rangle \tag{5}$$

(β_k is the normalization multiplier), but have well documented numerical stability issues in finite precision arithmetic with dimensions in excess of several hundred [13,20,27]. The root cause of the numerical stability problems appears to be twofold. First, Eq. (4) and, after iteration loop unrolling, Eq. (5) contain the Liouvillian raised to power n, meaning that the difference between the smallest and the largest $\hat{\hat{L}}^n$ eigenvalue increases exponentially during the iterations. Assuming (very conservatively) a factor-of-two spread between the smallest and the biggest eigenvalues of $\hat{\hat{L}}$ (e.g. ω and 2ω in a two-spin system), the difference would scale as 2^n with the number of iterations, meaning that the double-precision machine arithmetic will be overwhelmed around $n = 64$ and severe numerical rounding errors will creep in. Second, the Gram-Schmidt orthogonalization step implicit in Eq. (4) and explicitly applied to Eq. (5) is known to be numerically unstable and quite expensive for large values of n [28]. In other words, Lanczos pruning works well for small values of n (in practice up to about 500). Numerically stable orthogonalization techniques (based on SVD and QR factorization) that can deal with bigger state spaces [29] are computationally expensive and difficult to parallelize. The Liouvillian can, in principle, be preconditioned to have unit eigenvalues, but in practice performing such a precise preconditioning is equivalent to diagonalization [30]. The exponential propagator in the original Krylov subspace in Eq. (3) is naturally free of this problem, because the eigenvalues have a modulus of one, but computing the matrix exponential is equivalent, from the asymptotic scaling point of view, to diagonalization [25,26].

3. Zero track elimination

As we saw above, mapping Krylov subspaces for either the Liouvillian or the propagator is computationally expensive (at least quadratic and sometimes cubic scaling with the matrix size) and may generate numerical accuracy problems. We can, however, try to reformulate the original question—instead of looking for the vectors that *do* appear in the Krylov subspace and then projecting into that subspace, we can look for the vectors that *do not* appear there and project them out. It turns out that for spin dynamics simulations this procedure yields an efficient and numerically stable technique that we called Zero Track Elimination (ZTE).

It is well known that the typical initial state vectors in magnetic resonance simulations ($\hat{S}_Z, \hat{S}_+$ etc.) are very sparse. This sparsity is entirely separate from the sparsity of the Liouvillian—the latter refers to the general properties and connectivities of the spin system, whereas the former is a property of the actual pulse sequence or experiment through which the system evolves. It is also common wisdom that in a $\{\hat{\rho}_0, \hat{\hat{P}}\hat{\rho}_0, \hat{\hat{P}}^2\hat{\rho}_0, \ldots\}$ sequence many elements of the state vector $\hat{\rho}_0$ will, in fact, stay zero throughout the calculation, exactly or approximately. It seems likely that the dimensions corresponding to those elements may be pruned out. In fact, the following theorem holds:

Theorem 1 (Zero Track Theorem). *If a state-space vector $|i\rangle$ does not appear in the system density matrix anywhere within the first Larmor time step Δt of the evolution under a constant Liouvillian $\hat{\hat{L}}$, it will never appear in any of the subsequent steps.*

$$\langle i|e^{-i\hat{\hat{L}}t}|\hat{\rho}_0\rangle = 0, t \in [0, \Delta t] \;\Rightarrow\; \langle i|e^{-i\hat{\hat{L}}t}|\hat{\rho}_0\rangle = 0,\; t \in [0, \infty) \tag{6}$$

Proof. because the Taylor expansion of a continuous function is unique in any given interval, the following series, in order to be zero everywhere within $t \in [0,\Delta t]$, must have zero coefficients:

$$\begin{aligned}\langle i|e^{-i\hat{\hat{L}}t}|\hat{\rho}_0\rangle \;&= \langle i|\hat{\hat{E}}|\hat{\rho}_0\rangle + \langle i|\hat{\hat{L}}|\hat{\rho}_0\rangle(-it) + \\ &\langle i|\hat{\hat{L}}^2|\hat{\rho}_0\rangle\tfrac{(-it)^2}{2!} + \langle i|\hat{\hat{L}}^3|\hat{\rho}_0\rangle\tfrac{(-it)^3}{3!} + \cdots\end{aligned} \tag{7}$$

meaning that

$$\langle i|\hat{\hat{E}}|\hat{\rho}_0\rangle = \langle i|\hat{\hat{L}}|\hat{\rho}_0\rangle = \langle i|\hat{\hat{L}}^2|\hat{\rho}_0\rangle = \cdots = 0 \tag{8}$$

or, in other words, that $|i\rangle$ is orthogonal to the Krylov subspace generated by $\hat{\hat{L}}$ and $|\hat{\rho}_0\rangle$. The fact that the coefficients in (8) are zero means that the series in Eq. (7) will stay zero for all values of t. □

As may be seen from the proof, the ZTE pruning amounts to detecting the states that do not appear in the Krylov subspace generated by the Liouvillian and the initial state and pruning them out. This strategy is, in effect, the reverse of that employed by Lanczos pruning procedure, where the Krylov subspace is first mapped in its entirety and then projected into. The problem of actually computing the first time propagation step (or a number of steps, to avoid accidental zeros) can be circumvented by avoiding the explicit calculation of matrix exponentials. A computationally efficient procedure (itself based on Krylov subspaces) exists [26,31] for computing $e^{-i\hat{\hat{L}}t}|\hat{\rho}_0\rangle$ directly from $\hat{\hat{L}}$ and $|\hat{\rho}_0\rangle$, and the program code listed in the Supplementary Information makes use of it.

While Theorem 1 does provide a good start, the number of scalar products $\langle i|e^{-i\hat{\hat{L}}t}|\hat{\rho}_0\rangle$ that would stay *identically* zero is usually quite small. A much larger class will stay *approximately* zero and we need a robust pruning criterion for those, as well as an error estimate. We would ideally like to compute some small number of Δt steps and prune out the dimensions corresponding to the state vectors, whose magnitude had remained below a certain threshold. We will now demonstrate that this procedure would amount to ignoring the contributions from high-order vectors of the Krylov sequence generated by $\hat{\hat{L}}$ and $|\hat{\rho}_0\rangle$.

Theorem 2 (Thin Track Theorem). *If the absolute value of a scalar product $\langle i|e^{-i\hat{\hat{L}}t}|\hat{\rho}_0\rangle$ stays smaller than a fixed constant number $\varepsilon \ll 1$ throughout the first Larmor time step of the simulation, removing $\langle i|$ from the basis set is equivalent to restricting the simulation to a Krylov subspace generated by $\hat{\hat{L}}$ and $|\hat{\rho}_0\rangle$, ignoring the contributions to that subspace from high values of n in $\hat{\hat{L}}^n|\rho_0\rangle$.*

Proof. we observe that for a sub-Larmor time step Δt, the Taylor series for the scalar product $\langle i|e^{-i\hat{\hat{L}}t}|\hat{\rho}_0\rangle$ converges monotonically (that is, the absolute value of each subsequent term is smaller than that of the preceding one). Since the absolute value of the entire series is bounded from above by ε,

$$\langle i|e^{-i\hat{\hat{L}}t}|\hat{\rho}_0\rangle = \left|\sum_n \frac{(-it)^n}{n!}\langle i|\hat{\hat{L}}^n|\hat{\rho}_0\rangle\right| < \varepsilon \tag{9}$$

the high-order terms in this expansion are necessarily much smaller than ε for all values of t. That is, to within some small tolerance, the contribution to the propagator from $\hat{\hat{L}}^n|\rho_0\rangle$ does not contain any $|i\rangle$, which is therefore absent from the Krylov subspace of interest. Excluding this state (and others that fall below the ε threshold) from the basis set would therefore amount to restricting the

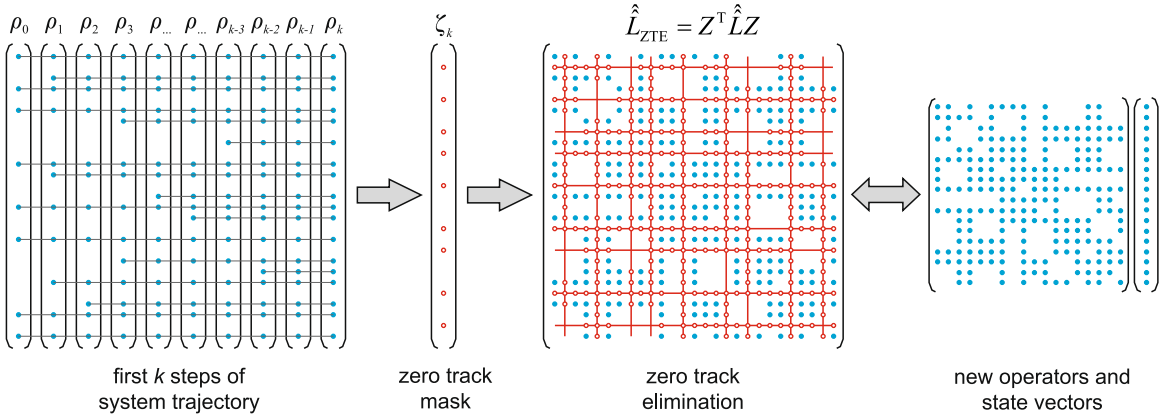


Fig. 3. Zero Track Elimination algorithm schematic. The initial density matrix ρ_0 (shown as a vector with blue dots denoting non-zeros) is propagated forward in time for k Larmor steps using Krylov technique [26,31] to avoid the computation of Liouvillian exponential. Zero tracks in the density matrix are detected and the zero mask ζ_k (red dots denote zero tracks) is generated. Eq. (11) is then applied: the rows and columns that would operate on zero tracks are erased from the Liouvillian and the elements corresponding to zero tracks are erased from the density matrix. A return to the full state space may be performed at any time in the subsequent simulation by re-inserting zeroes into their original positions within the density matrix and returning to the original Liouvillian. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

simulation to the Krylov subspace in Eq. (4) and ignoring the contributions from the high-order terms. □

Because ZTE restricts the system to low-order vectors of the Krylov subspace, it accomplishes the same goal as Lanczos pruning. It does, however, turn out to be faster, because scalar products in Eq. (9) can be replaced by a simple zero check within the density matrix.[3] The practical implementation of ZTE technique (illustrated schematically in Fig. 3, the Matlab source code is listed in the Supplementary Information) contains the following stages:

1. Generation of the fist k trajectory steps of approximately Larmor size using Krylov propagation [26,31] to avoid computing the full Liouvillian exponential. The procedure scales as vector-matrix multiplication, that is, between $O(N^2)$ and $O(N)$, depending on sparsity, with the dimension of the Liouvillian matrix.
2. Detection of zero tracks in the resulting trajectory, which proceeds by a simple tolerance check, and generation of the zero track mask ζ.

$$\zeta_i = \begin{cases} 1 & \text{if } \sum_k |\rho_i^{(k)}| < \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

3. Construction of the shrinking matrix Z, which is a unit matrix with columns flagged in the zero track mask ζ taken out. It is easy to see that the following transformation would then accomplish the shrinkage

$$\hat{\hat{L}}_{\mathrm{ZTE}} = Z^{\mathrm{T}}\hat{\hat{L}}Z \quad \hat{\rho}_{0,\mathrm{ZTE}} = Z^{\mathrm{T}}\hat{\rho}_0 \tag{11}$$

   Eq. (11) is a formality—in practical calculations the computer is simply instructed to throw a number of rows and columns away from $\hat{\hat{L}}$ and $|\hat{\rho}_0\rangle$.
4. Calculation of the rest of the trajectory in the resulting reduced state space using standard spin dynamics simulation techniques. The reduced operator matrices and state vectors constitute a low-dimensional *representation* of the system dynamics, meaning that all the standard simulation techniques (exponential propagation, diagonalization, scalar products etc.) work as before, making the procedure easy to add to any existing simulation code.
5. Optionally, a return to the full state space by re-inserting zeros into their original positions in the state vector and returning to the original Liouvillian. This may be necessary if the next pulse sequence segment has a different Liouvillian (the detailed treatment of time-dependent cases is given in a separate section below).

The procedure described above can be made more general (but more computationally expensive) if the explicit zero check is replaced by scalar product check prescribed by Theorems 1 and 2. Interestingly, because the index of the pruned density matrix dimensions can be kept, they can be re-introduced as zeros at any time, meaning that the procedure is reversible. The speed advantage over Lanczos and SVD techniques comes from the lack of actual matrix-vector multiplication in stages 2 and 3. All operations amount to array element removal.

4. ZTE with time-dependent Liouvillians

The discussion has so far referred implicitly to time-independent Liouvillians. We will now generalize the ZTE formalism and demonstrate that it remains valid and computationally efficient in the two common time-dependent cases: for piecewise-constant and continuously varying Liouvillians.

4.1. Piecewise-constant Liouvillians

Many pulse sequences in NMR and EPR can be approximated as a series of hard pulses and delays with different (due to decoupling, gradients *etc.*) static Liouvillians, meaning that spin evolution under a piecewise-constant Liouvillian has to be considered. A switch between static Liouvillians triggers a jump to a different Krylov subspace in Eq. (4), requiring a brief return to the full state space and a repetition of the ZTE procedure with the new Liouvillian. This move to a different Krylov subspace is shown schematically in Fig. 1 as a subspace jump by the system trajectory. The implementation schematic (Matlab source code is included in the Supplementary Information) is given in Fig. 4. At the start of each Liouvillian segment, ZTE pruning is applied and an index of the zero tracks (a list of their original positions in the state vector) is kept. The simulation is carried out in the reduced state space for

[3] The reason for this is highly non-trivial and probably specific to spin dynamics—the basis set is usually generated with a product operator procedure, starting from Pauli matrices. Because the Liouvillian is generated in the same way, it takes a particularly sparse and simple form.

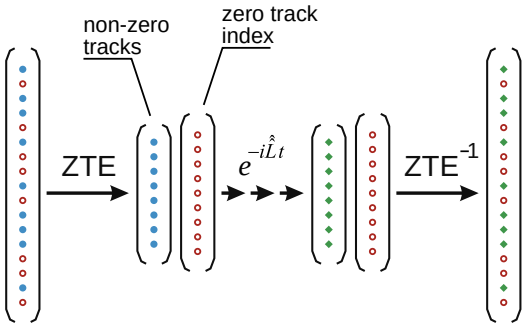


Fig. 4. ZTE algorithm schematic for a piecewise-continuous Liouvillian. For each Liouvillian segment, ZTE pruning is applied, the temporal propagation is carried out in the reduced state space and then a return to the full state space is performed. For further details, see text.

the prescribed length of time and a return to the full state space is performed by re-inserting the zeros into their original positions. The resulting full state-space vector is used as a starting state for the next Liouvillian segment.

4.2. Continuously varying Liouvillians

Any continuously time-dependent Liouvillian operator admits the following expansion

$$\hat{\hat{L}}(t) = \sum_i f_i(t)\hat{\hat{F}}_i, \tag{12}$$

where $f_i(t)$ are linearly independent continuous functions of time, serving as expansion coefficients for $\hat{\hat{L}}(t)$ in a static operator basis $\{\hat{\hat{F}}_i\}$. In the simplest case of scalar time dependence of the form $\hat{\hat{L}}(t) = f(t)\hat{\hat{F}}$ (e.g. a Gaussian pulse, where only the amplitude is varied) both Theorems 1 and 2 are still valid in full generality because the envelope function $f(t)$ only re-scales the time step

$$\exp(-i\hat{\hat{L}}(t)\Delta t) = \exp(-i\hat{\hat{F}}f(t)\Delta t) = \exp(-i\hat{\hat{F}}\Delta t') \quad \Delta t' = f(t)\Delta t \tag{13}$$

which does not affect zero tracks—they stay zero because no extra coherence transfer paths are introduced by the scalar modulation. Thus, the ZTE pruning only needs to be performed once at the start of the simulation with, say, $f(t) = 1$. If several independently modulated components are present in Eq. (12), the time step would be re-scaled differently for every component

$$\exp[-i\hat{\hat{L}}(t)\Delta t] = \exp\left[-i\left(\sum_i f_i(t)\hat{\hat{F}}_i\right)\Delta t\right] = \exp[-i(\hat{\hat{F}}_1\Delta t' + \hat{\hat{F}}_2\Delta t'' + \hat{\hat{F}}_3\Delta t''' + \ldots)]. \tag{14}$$

If the time steps $\{\Delta t', \Delta t'', \Delta t''', \ldots\}$ are independent and non-zero within the ZTE zero testing interval, their subsequent time dependence would not introduce any new coherence transfer pathways—it would only re-balance and possibly remove the contributions from the existing ones. Any zero tracks identified within such testing interval would therefore stay zero. This reasoning can be summarized by a simple qualitative rule: *if a representative segment of the Liouvillian time dependence is used in the ZTE procedure, ZTE is valid for the time-dependent case and only needs to be performed once*. This rule explains why piecewise-constant Liouvillians were treated as a separate case above—individual time segments of a piecewise-constant matrix function tend not to be representative of each other.

5. Simulation results and conclusions

Fig. 5 demonstrates the performance of Zero Track Elimination and Lanczos pruning. A fairly typical spectrum of a strongly coupled 10-spin system (Fig. 5C) requires 1,048,576 states for the exact Liouville-space simulation, 895 states after the SSR(4) procedure [17] and a mere 280 states after ZTE or Lanczos pruning had been applied. The bottom spectrum in Fig. 5C illustrates the numerical instability exhibited by Lanczos pruning (even with SVD orthogonalization) with large state spaces. Zero Track Elimination procedure avoids computing high powers of the Liouvillian and is numerically stable.

Lanczos pruning, however, tends to perform better for small, densely coupled spin systems and tighter state-space restriction. For a fully coupled (every spin to every other spin) five-spin system in very small state spaces (Fig. 5A and B), the Lanczos pruning reproduces the spectral envelope appreciably better than ZTE. This can be important in solid-state and EPR simulations, where powder averages are much more sensitive to the spectral envelopes of individual orientations than to the accuracy of the splittings inside

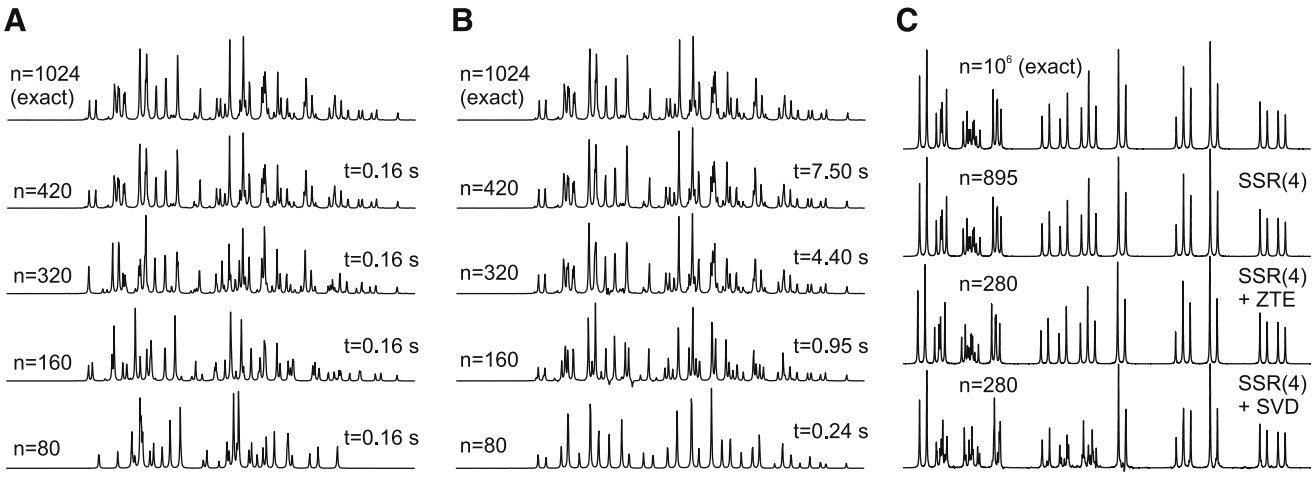


Fig. 5. Zero Track Elimination (A) and Lanczos pruning (B) performance for a 90°-acquire experiment on a fully and strongly coupled (every spin to every other spin, with scalar coupling amplitudes selected randomly within 0–20 Hz interval) system of five spins 1/2. The dimension of the reduced state space is given on the left of each column and the CPU time (Opteron 265/ DDR400) spent in pruning is shown on the right. The simulations labeled 'exact' were performed in the complete state space. With larger spin systems (a 90°-acquire experiment simulation on a system with 10 spins 1/2 is shown in panel C, see the source code in the Supplementary Information for the simulation details and the spin system structure), both techniques can be applied on top of SSR (restriction to four-spin orders between directly connected spins [17] is used here), resulting in accurate simulations with very small matrices.

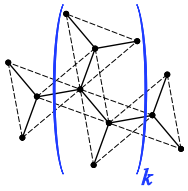
Table 1
State-space dimension and matrix density statistics for the exact, SSR(4) and SSR(4) + ZTE simulation of a 90°-acquire experiment in large coupled spin systems

Number of spins[a]	State-space dimension			SSR(4) Liouvillian density, %[c]	SSR(4) + ZTE Liouvillian density, %[c]
	Full	SSR(4)[b]	SSR(4) + ZTE		
6	4096	1747	212	0.63	4.0
12	1.7×10^7	15913	577	0.068	1.4
18	6.9×10^{10}	37981	655	0.028	1.1
24	2.8×10^{14}	60985	795	0.017	1.0
54	3.2×10^{32}	1.7×10^5	3418	5.7×10^{-3}	0.18
102	2.6×10^{61}	3.6×10^5	7874	2.8×10^{-3}	0.08
198	1.6×10^{119}	7.3×10^5	14282	1.4×10^{-3}	0.05

[a] See Eq. (15) for the coupling structure of the model spin systems.
[b] SSR(4): state-space restriction up to four-spin orders between directly connected spins [17].
[c] Percentage of non-zero elements in the matrix.

them. Because the orthogonalization step is computationally expensive, Lanczos pruning tends to become more expensive as well, as the restricted space gets bigger. On a side note, and somewhat counterintuitively, even in this *fully* coupled spin system, more than half of the spin states can still be ignored.

Simulation results for larger spin systems are given in Table 1. At least for the simple pulse-acquire experiment, ZTE procedure cuts the state-space dimensions by further two orders of magnitude and makes the operator matrices quite dense, suggesting that the pruning is now complete. The coupling patterns in the test systems in Table 1 are representative of a typical pattern of *J*-couplings found in proteins:



(15)

where solid lines correspond to a *J*-coupling randomly selected from a 0-50 Hz interval and the dashed lines indicate a random *J*-coupling within a 0–10 Hz interval. As Table 1 demonstrates, the combined SSR(4) + ZTE (*i.e.* state-space restriction up to four-spin orders between directly connected spins followed by zero track elimination) procedure reduces the dimension of the matrices involved in the simulation by many orders of magnitude, cutting it from astronomical to manageable. It is also worth noting that, while large spin systems with regular coupling patterns (e.g. linear chains [14–16]) have been treated in the literature, the algorithms outlined above and in Ref. [17] differ in that they can deal with arbitrary scalar or tensor coupling patterns.

In summary, we believe that the combination of the state-space restriction procedure [17] with Zero Track Elimination outlined above and/or Lanczos pruning [13,20] constitutes a computationally efficient and numerically stable Minimal Truncation Scheme [13,20] for spin dynamics simulations. It appears that the real matrix dimensions required for simulations are, in fact, tiny, and the astronomically sized matrices generated by the traditional direct product procedures are completely unnecessary.

Acknowledgments

The author is grateful to Prof. Jack Freed, Prof. Peter Hore, Dr. Paul Hodgkinson and Dr. Stefan Stoll for stimulating discussions. This work is supported by EPSRC Grant EP/F065205/1.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmr.2008.08.008.

References

[1] M. Bak, J.T. Rasmussen, N.C. Nielsen, SIMPSON: a general simulation program for solid-state NMR spectroscopy, J. Magn. Reson. 147 (2000) 296–330.
[2] R.R. Ernst, G. Bodenhausen, A. Wokaun, Principles of Nuclear Magnetic Resonance in One and Two Dimensions, Clarendon Press, Oxford, 1987.
[3] P. Hodgkinson, L. Emsley, Numerical simulation of solid-state NMR experiments, Progr. NMR Spec. 36 (2000) 201–239.
[4] M. Mehring, V.A. Weberruss, Object-oriented magnetic resonance: classes and objects, calculations and computations, Academic Press, San Diego, California, 2001.
[5] N.C. Nielsen, Computational aspects of biological solid-state NMR spectroscopy, in: A. Ramamoorthy (Ed.), NMR Spectroscopy of Biological Solids, CRC/ Taylor & Francis, 2006. pp. 261–304.
[6] A.C. Sivertsen, M. Bjerring, C.T. Kehlet, T. Vosegaard, N.C. Nielsen, Numerical simulations in biological solid-state NMR spectroscopy, Annu. Rep. NMR Spectr. 54 (2005) 243–293.
[7] S.A. Smith, T.O. Levante, B.H. Meier, R.R. Ernst, Computer simulations in magnetic resonance. An object-oriented programming approach, J. Magn. Reson. 106 (1994) 75–105.
[8] M. Veshtort, R.G. Griffin, SPINEVOLUTION: a powerful tool for the simulation of solid and liquid state NMR experiments, J. Magn. Reson. 178 (2006) 248–282.
[9] A. Jerschow, MathNMR: spin and spatial tensor manipulations in mathematica, J. Magn. Reson. 176 (2005) 7–14.
[10] I. Kuprov, N. Wagner-Rundell, P.J. Hore, Bloch–Redfield–Wangsness theory engine implementation using symbolic processing software, J. Magn. Reson. 184 (2007) 196–206.
[11] R.S. Dumont, S. Jain, A. Bain, Simulation of many-spin system dynamics via sparse matrix methodology, J. Chem. Phys. 106 (1997) 5928–5936.
[12] P. Hodgkinson, D. Sakellariou, L. Emsley, Simulation of extended periodic systems of nuclear spins, Chem. Phys. Lett. 326 (2000) 515–522.
[13] G. Moro, J.H. Freed, Calculation of ESR spectra and related Fokker-Planck forms by the use of the Lanczos algorithm, J. Chem. Phys. 74 (1981) 3757–3773.
[14] S.I. Doronin, E.B. Fel'dman, E.I. Kuznetsova, G.B. Furman, S.D. Goren, Dipolar temperature and multiple-quantum NMR dynamics in dipolar ordered-spin systems, JETP Lett. 86 (2007) 24–27.
[15] S.I. Doronin, E.B. Fel'dman, S. Lacelle, Multiple-quantum nuclear magnetic resonance spin dynamics in disordered rigid chains and rings, J. Chem. Phys. 117 (2002) 9646–9650.
[16] E.B. Fel'dman, R. Bruschweiler, R.R. Ernst, From regular to erratic quantum dynamics in long spin 1/2 chains with an XY Hamiltonian, Chem. Phys. Lett. 294 (1998) 297–304.
[17] I. Kuprov, N. Wagner-Rundell, P.J. Hore, Polynomially scaling spin dynamics simulation algorithm based on adaptive state-space restriction, J. Magn. Reson. 189 (2007) 241–250.
[18] Z. Bai, Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems, Appl. Numer. Math. 43 (2002) 9–44.
[19] R.W. Freund, Model reduction methods based on Krylov subspaces, Acta Numer. 12 (2003) 267–319.
[20] K.V. Vasavada, D.J. Schneider, J.H. Freed, Calculation of ESR spectra and related Fokker-Planck forms by the use of the Lanczos algorithm. II. Criteria for truncation of basis sets and recursive steps utilizing conjugate gradients, J. Chem. Phys. 86 (1986) 647–661.
[21] N.C. Nielsen, O.W. Sorensen, Accessible states in Liouville space. A two-dimensional extension of the universal bound on spin dynamics applied to polarization transfer in INSM Spin- systems, J. Magn. Reson. 99 (1992) 449–465.
[22] N.C. Nielsen, O.W. Sorensen, 2D bounds on polarization transfer involving quadrupolar spin nuclei, J. Magn. Reson. 99 (1992) 214–222.
[23] J.D. Van Beek, M. Carravetta, G.C. Antonioli, M.H. Levitt, Spherical tensor analysis of nuclear magnetic resonance signals, J. Chem. Phys. 122 (2005) 1–12.
[24] P. Guntert, N. Schaefer, G. Otting, K. Wuthrich, POMA: a complete mathematica implementation of the NMR product-operator formalism, J. Magn. Reson. A 101 (1993) 103–105.

[25] C. Moler, C. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later, SIAM Rev. 45 (2003) 3–49.
[26] R.B. Sidje, Expokit: a software package for computing matrix exponentials, ACM Trans. Math. Soft. 24 (1998) 130–156.
[27] E. Cahill, A. Irving, C. Johnston, J. Sexton, Numerical stability of Lanczos methods, Nucl. Phys. B 83-84 (2000) 825–827.
[28] N.J. Higham, Accuracy and Stability of Numerical Algorithms, Society for Industrial and Applied Mathematics, Philadelphia, 1996.
[29] S. Gugercin, An iterative SVD-Krylov based method for model reduction of large-scale dynamical systems, Lin. Alg. Appl. 428 (2008) 1964–1986.
[30] A.V. Knyazev, Preconditioned eigensolvers—An oxymoron?, Electr Trans. Num. Anal. 7 (1998) 104–123.
[31] M. Hochbruck, C. Lubich, On Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal. 34 (1997) 1911–1925.