Inverse-covariance matrix of linear operators for quantum spectrum scanning

Hamse Y. Mussa[1], Jonathan Tennyson[2], and Robert C. Glen[1]

[1]Unilever Centre for Molecular Sciences Informatics
Department of Chemistry, University of Cambridge
Lensfield Road, Cambridge CB2 1EW, UK
[2]Department of Physics and Astronomy
University College London, London WC1E 6BT, UK

E-mail: hym21@cam.ac.uk

Abstract. It is demonstrated that the Schrödinger operator in $\hat{\mathbf{H}} \mid \psi_k >= E_k \mid \psi_k >$ can be associated with a covariance matrix whose eigenvalues are the squares of the spectrum $\sigma(\hat{\mathbf{H}} + \mathbf{I}\zeta)$ where ζ is an arbitrarily chosen shift. An efficient method for extracting $\sigma(\hat{\mathbf{H}})$ components, in the vicinity of ζ, from a few specially selected eigenvectors of the inverse of the covariance matrix is derived. The method encapsulates (and improves on) the three most successful quantum spectrum scanning schemes: Filter-Diagonalization, Shift-and-invert Lanczos and Folded Spectrum Method. It gives physical insight into the scanning process. The new method can also be employed to probe the nature of underlying potential energy surfaces. A sample application to the near-dissociation vibrational spectrum of the HOCl molecule is presented.

PACS numbers: 03.65.-w,02.70.-c,02.10.10.Yn,02.30.Tb,02.50.Sk,02.30.Zz,02.30Yy

Keywords: Eigensolver, Mixed state, Quantum Control

1. Introduction

Quantum mechanics provides our understanding of the behaviour of microscopic objects such as atoms, molecules and their constituents. Thus non-relativistic studies of these objects and any process involving them essentially requires solving the appropriate Schrödinger equation. However, apart from few cases, solving this equation accurately is a major problem — in fact it was identified as a computational challenges of the last century and still remains one [1]. A well documented examples which reflect this fact are the difficulties one faces in order to understand/explain chemical reactions, molecular spectroscopy, and thermodynamic properties from first principles [1, 2, 3, 4].

Usually, the time-independent Schrödinger eigenproblem

$$\hat{\mathbf{H}} \mid \psi_k >= E_k \mid \psi_k >, \tag{1}$$

of the Hamiltonian operator $\hat{\mathbf{H}}$ is solved for the eigenpairs $(E_k,\mid \psi_k >)$ which are the quantum mechanically allowed energy values and their associated eigenstates of the physical system. In this work we assume that the eigenvalue spectrum is discrete and finite: $k = 1, 2, ..., n$. Generally $\hat{\mathbf{H}}$ is real symmetric or Hermitian. However, non-Hermiticity can arise, in particular, when studying dissipative quantum states which often play an important role in energy transfer processes, such as scattering and unimolecular reactions, see [4, 5] and references therein.

The eigenvalue problem is typically handled by expanding $\mid \psi_k >$ in an appropriately chosen basis set $\{\mid \phi_j >\}_{j=1}^{n}$, $\mid \psi_k >= \sum_{j=1}^{n} u_{jk} \mid \phi_j >$, and then finding the eigenpairs of $\hat{\mathbf{H}}$ in a n- dimensional Hilbert space spanned by the set. For clarity and without loss of generality, we will assume that $\hat{\mathbf{H}}$ is real symmetric; generalisation to non-Hermitian operators is straightforward.

Since in finite-dimensional Hilbert spaces all operations, physical or otherwise, simplify to manipulating ordinary linear algebra, finding solutions of Eq. (1) amounts to solving an eigenvector problem

$$\acute{\mathbf{H}}\mathbf{u}_k = E_k \mathbf{u}_k \tag{2}$$

In other words diagonalising $\acute{\mathbf{H}} = (< \phi_i \mid \hat{\mathbf{H}} \mid \phi_j >)_{i,j=1}^{n}$, a n-dimensional matrix representation of $\hat{\mathbf{H}}$ in $\{\mid \phi_j >\}_{j=1}^{n}$, for yielding E_k and its corresponding eigenvector $\mathbf{u}_k \in \mathcal{R}^{n\times 1}$ which is the $\mid \psi_k >$ representative in the chosen basis set — that is, $\mathbf{u}_k$ elements are the values of different features which characterise $\mid \psi_k >$ in the space defined by the basis set.

With conventional matrix diagonalisers one has to calculate the entire spectrum and the corresponding eigenvectors of $\acute{\mathbf{H}}$, from the bottom up even when one wishes to examine only eigenstates with specific energy levels [2]. Furthermore the computational work with canonical eigensolvers scale badly with n, as n^3. Thus, with the current computer power, these diagonalisers are useful only for studying physical systems represented by small and moderately large matrices.

The last two decades have seen the development of methods that can scan the energy spectrum region-by-region as desired by employing function operators: A commonly

used function operator is the Green operator $(\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1}$, where ζ is the centre of the energy region that one wishes to scan, but ζ itself is not an eigenvalue of $\acute{\mathbf{H}}$.

The most successful of these methods are filter-diagonalization (FD) [3, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19], Shift-and-invert Lanczos (SaiL) [2, 20, 21, 22, 23, 24] and Folded Spectrum Method (FSM) [25, 26]. Despite their remarkable success these algorithms have shortcomings. Here we propose and derive a method which not only improves them, but also encapsulates all three methods.

The new method can also be employed as a probe of the nature of potential energy surfaces, which can give physical insight into system. We show that the Green operator $(\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1}$ is not a mere mathematical trick in the scanning schemes, but can be associated with the covariance matrix of a mixed quantum state composed of an ensemble of pure states. We also illustrate that the mixed state is related to what is commonly known as filtered states in the spectrum scanning literature. A possible connection between the proposed scanning scheme and quantum control theory [29] is also pointed out. All of this can be deduced from an analysis of $(\acute{\mathbf{H}} - \mathbf{I}\zeta)^2$.

The following section gives a brief pedagogical introduction to FD, SaiL and FSM using a simple framework with which all the scanning schemes can be analysed. The framework lacks mathematical rigor, but is adequate for the intended purpose. In each case a complete treatment of the methods is given in the cited references. The proposed Covariance Based Eigensolver (CoBaE) algorithm is derived in Section 3 and some of its implications are given in Section 4. Test results and their analyses are presented in Section 5. Section 6 gives our concluding remarks.

2. Overview of Quantum spectral scanning schemes

From the outset we note that scanning algorithms, including the proposed method, require a few n-dimensional vectors to be stored in computer memory at one time. This means that the algorithms are not universal — they do not work for problems where n is too huge for the few vectors to be kept in computer memory. In this scenario, currently there is no known algorithms which can accurately calculate $\sigma(\acute{\mathbf{H}})$. However, schemes, such the stochastic diagonalisers [27, 28], can yield approximate eigenpairs.

We develop the framework from the following basic notions in canonical quantum formalism.

(i) An isolated quantum system is described in terms of abstract vectors and operators in a Hilbert space $\mathcal{H}$. The state vectors represent quantum mechanically allowable states of the system while its physical observables are associated with the operators whose actions determine how the state vector resides in $\mathcal{H}$.

(ii) Hamiltonian operators can be both Hermitian and unitary (real-symmetric and orthonormal) if and only if the modulus of E_k is 1, $\forall k$, *i.e.*, $E_k = \pm 1$. Most unitary operators are not Hermitian and therefore cannot represent physical observables. Nevertheless there are some Hermitian-unitary operators, fortunately their corresponding matrices can be diaganalised quite easily. The majority

of Hamiltonian quantum systems are associated with matrices that are real symmetric/Hermitian, $\acute{\mathbf{H}} = \acute{\mathbf{H}}^T$, but not orthonormal/unitary, $\acute{\mathbf{H}}^T\acute{\mathbf{H}} \neq \mathbf{I}$, and difficult to solve.

(iii) A Hilbert space $\mathcal{H}$ can be spanned by infinitely many suitable basis sets. For the finite Hilbert space of dimension n, one possible basis set consists of all the eigenvectors of $\acute{\mathbf{H}}$, $\{\mathbf{u}_k\}_{k=1}^n$. Due to the linearity of $\mathcal{H}$, any possible state vector $\mathbf{b}$ can be expressed as $\mathbf{b} = \sum_{k=1}^n c_k\mathbf{u}_k$.

$\acute{\mathbf{H}}^T\acute{\mathbf{H}} \neq \mathbf{I}$ suggests that $\mathbf{b}^T\mathbf{b} \neq (\acute{\mathbf{H}}\mathbf{b})^T(\acute{\mathbf{H}}\mathbf{b})$, *i.e.*, the norm of $\mathbf{b}$ is not conserved. The operation of $\acute{\mathbf{H}}$ upon $\mathbf{b}$ transforms $\mathbf{b}$ into another state vector in $\mathcal{H}$ in one of two ways: (a) The magnitude of $\mathbf{b}$ is dilated, but its direction is left unchanged, *i.e.*, $\mathbf{b}$ is an eigenvector of $\acute{\mathbf{H}}$; (b) One of the eigenvectors in the expansion $\mathbf{b} = \sum_{k=1}^n c_k\mathbf{u}_k$ is magnified/stretched more than the others. In this case $\mathbf{b}$ is skewed towards the eigenvector whose length increased the most, the so-called dominant eigenvector $\mathbf{u}_d$.

Linearly shifting $\acute{\mathbf{H}}$ by ζ gives $(\acute{\mathbf{H}} - \zeta\mathbf{I})$. Elementary linear algebra [30, 31, 32, 33] shows that the action of $(\acute{\mathbf{H}} - \zeta\mathbf{I})^\kappa$, where κ is a positive or negative integer, on $\mathbf{b}$ shrinks/folds or respectively dilates/magnifies the eigenvectors in the expansion $\mathbf{b} = \sum_{k=1}^n c_k\mathbf{u}_k$ whose corresponding eigenvalues are near ζ. The eigenvector which is folded the most is the smallest eigenvector, $\mathbf{u}_s$, of $(\acute{\mathbf{H}} - \zeta\mathbf{I})^{\kappa>0}$, whereas the eigenvector that is magnified the most is $\mathbf{u}_d$ of $(\acute{\mathbf{H}} - \zeta\mathbf{I})^{\kappa<0}$. Henceforward the terms 'smallest/largest eigenvalue' and 'smallest/largest eigenvector' mean the eigenvalue that is the smallest/largest in magnitude and the eigenvector whose corresponding eigenvalue is the smallest/largest in the sense aforementioned.

The discussion above can be generalised to any function operator of the Hamiltonian system, $f(\acute{\mathbf{H}}|\zeta)$, that satisfies the following: (a) $\forall\zeta$ in $[E_1, E_n]$, provided that $\zeta \neq E_k$, and $f(E_k \mid \zeta))$ exists and is finite; (b) $(\mathbf{u}_k, f(E_k \mid \zeta)) \Rightarrow (E_k, \mathbf{u}_k)$, an eigenpair of $\acute{\mathbf{H}}$ [34, 35]. However, due to perhaps the fundamental role that it plays in many physical applications, the Green operator $(\acute{\mathbf{H}} - \zeta\mathbf{I})^{-1}$ is the most commonly employed function operator in the majority of spectrum scanning approaches [2]. It is this operator and those closely related to it, in particular $(\acute{\mathbf{H}} - \zeta\mathbf{I})^p$ with $p = 1, -2, 2$, we are principally concerned with.

Now we come to the central idea of the framework. If one wishes to calculate $\mathbf{u}_d$ of a function operator (say, a magnifier), a method as simple as the power method should suffice provided that the operation $(\acute{\mathbf{H}} - \zeta\mathbf{I})^p\mathbf{b}$ can be performed [30, 31, 32, 33]. Clearly it would be highly desirable to extract a portion of the energy spectrum rather than one eigenpair at a time. Sophisticated subspace based methods [30, 33] are suitable for this objective.

It can be surmised from (iii) above that any subset of $\{\mathbf{u}_k\}_{k=1}^n$, say $\{\mathbf{u}_{k+j}\}_{j=1}^l (l \ll n)$, determines a l-dimensional subspace $\mathcal{S} \subset \mathcal{H}$ where the corresponding eigenvalues lie in the energy range $[E_{k+1}, E_{k+l}]$ [30, 31]. Like $\mathcal{H}$, $\mathcal{S}$ can also be spanned by other basis sets. Let $\{\mathbf{s}_i\}_{i=1}^l$, $\mathbf{s}_i \in \mathcal{R}^{n\times 1}$, be one such basis set and for clarity sake assume that it is orthogonal. As the eigenvectors of $\acute{\mathbf{H}}$ in $\mathcal{S}$ are expressible in terms of $\{\mathbf{s}_i\}_{i=1}^l$:

$\mathbf{u}_{k+j} = \sum_{i=1}^{l} q'_{ji}\mathbf{s}_i$, the coefficients $\{\mathbf{q}'_j\}_{j=1}^{l}$ are the eigenvectors of a small matrix $\mathbf{Q} \in \mathcal{R}^{l\times l} = (\mathbf{s}_i, \acute{\mathbf{H}}\mathbf{s}_{i'})_{i,i'=1}^{l}$, the representation of matrix $\acute{\mathbf{H}}$ in $\{\mathbf{s}_i\}_{i=1}^{l}$; $\{q_j\}_{j=1}^{l}$ are the corresponding eigenvalues of $\mathbf{Q}$ and of $\acute{\mathbf{H}}$ in the energy range $[E_{k+1}, E_{k+l}]$ of $\sigma(\acute{\mathbf{H}})$.

The question of how to find $\{\mathbf{s}_i\}_{i=1}^{l}$ in order to build $\mathbf{Q}$ is the core one for the scanning methods. However, these methods differ from each other in the manner they compute, or rather estimate, the set.

2.1. SaiL method: 'a set of Lanczos vectors' as $\{\mathbf{s}_i\}_{i=1}^{l}$

Repeated operations on $\mathbf{b}$, with any function operator, generates a sequence known as the Krylov sequence [31, 33]. $\{((\acute{\mathbf{H}} - \zeta\mathbf{I})^{-1})^i\mathbf{b}\}_{i=0}^{l-1}$ is the Krylov sequence of the Green operator, where $((\acute{\mathbf{H}} - \zeta\mathbf{I})^{-1})^{l-1}\mathbf{b}$ is a vector pointing to the same direction as $\mathbf{u}_d$ [33] of the operator.

The elements of the sequence are linearly independent [31] and orthogonalising them results in a set of Lanczos vectors [33, 36, 37] $\{\mathbf{v}_i\}_{i=0}^{l-1}$, a viable basis set for $\mathcal{S}$ [31, 33], in which $\acute{\mathbf{H}}$ can be expressed in order to produce $\mathbf{Q}$. As explained before, the construction of $\mathbf{Q}$ is what is being sought: This matrix is easy to diagonalise and its eigenpairs subsequently provide the eigenpairs of $\acute{\mathbf{H}}$ in the immediate vicinity of ζ.

The essence of the SaiL scheme is generating this small Krylov sequence efficiently. The scheme was first introduced by Ericsson and Ruhe (E&R) where the multiplications $\{((\acute{\mathbf{H}} - \zeta\mathbf{I})^{-1})^i\mathbf{b}\}_{i=0}^{l-1}$ were performed via LU decomposition of $(\acute{\mathbf{H}} - \mathbf{I}\zeta)$. For small and moderately large $(\acute{\mathbf{H}} - \mathbf{I}\zeta)$, LU factorization is not only computationally feasible, but quickly converges the Krylov sequence [17, 22, 24]. However, as n increases computer memory requirements makes E&R's approach prohibitive. This drawback can be circumvented by iteratively solving the following linear system

$$(\acute{\mathbf{H}} - \mathbf{I}\zeta)\mathbf{w}_{i+1} = \mathbf{w}_i \tag{3}$$

where $i = 0, 1,.., l$ -1 ; $\mathbf{w}_0 = \mathbf{b}$; $\mathbf{w}_1 = (\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1}\mathbf{b}$.

Wyatt [2] was probably the first to adapt this strategy for molecular spectral scanning. It has since been used by a number of researchers including Carrington and co-workers [21, 23]. Unfortunately, $(\acute{\mathbf{H}} - \mathbf{I}\zeta)$ in Eq. (3) is ill-conditioned. Thus preconditioning it first is essential before one can employ suitable accelerators [33, 38, 39]. Poirier and Carrington [23] reported a remarkable reduction of the value of l when they appropriately pre-conditioned $(\acute{\mathbf{H}} - \mathbf{I}\zeta)$. Note that l determines both the computer memory required for storing the Lanczos vectors and how many times the linear system in Eq. (3) has to be solved.

To our knowledge there is no systematic way in which one can build a suitable pre-conditioner for a given Hamiltonian system [39]. Clearly this is a major hindrance to the computational efficiency of the SaiL algorithm.

Iung and Leforestier [40], and Kono [41] have also used spectrum scanning schemes where non-Green function operators were employed to generate the Krylov sequence.

2.2. FD: 'filtered state vectors' as $\{\mathbf{s}_i\}_{i=1}^{l}$

A second suitable basis set for $\mathcal{S}$ is $\{\mathbf{y}(\zeta_j)\}_{j=1}^{l}$ where $\mathbf{y}(\zeta_j) = (\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1}\mathbf{b}$, and $\{\zeta_j\}_{j=1}^{l} \in [E_{k+1}, E_{k+l}]$, but $\zeta_j \neq E_{k+j}$. Superficially $\mathbf{y}(\zeta_j)$, the so-called filtered state, can be seen as the projection of $\mathbf{b}$ into $\mathcal{S}$ by $(\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1}$ [3, 7, 8, 10, 17]. In this scheme, the spectrum of the Hamiltonian system is scanned over an energy range unlike SaiL. Orthogonalising the set $\{\mathbf{y}(\zeta_j)\}_{j=1}^{l}$ results in $\{\mathbf{s}_i\}_{i=1}^{l}$, in which $\acute{\mathbf{H}}$ can be represented in order to obtain $\mathbf{Q}$. Again diagonalising $\mathbf{Q}$ leads to yielding the eigenpairs of $\acute{\mathbf{H}}$ in the energy interval $[E_{k+1}, E_{k+l}]$.

The method was introduced by Neuhauser [6]. In its original form, the filtered states are generated in the time domain and $\{\mathbf{y}(\zeta_j)\}_{j=1}^{l}$ is the Fourier transform. Full details of Neuhauser's approach can be found in Ref. [6].

The method has been adapted and modified by several authors, see [3, 7, 8, 9, 17] and references therein, by building the filtered states in the energy domain. However, there are several schemes for "projecting" out $\{\mathbf{y}(\zeta_j)\}_{j=1}^{l}$ from $\mathbf{b}$. Mandelshtam and Taylor [7], and Chen and Guo [8] construct the set $\{\mathbf{y}(\zeta_j)\}_{j=1}^{l}$ by expanding magnifying function operators, such as the Green operator, in some kind of polynomials, usually Chebyshev polynomials,

$$\mathbf{y}(\zeta_j) = \sum_{i}^{K} q_i(\zeta_j) T_i(\overline{\mathbf{H}})\mathbf{b}. \tag{4}$$

where K is the order of expansion; $q_i(\zeta_j)$s are polynomial co-efficients; $\overline{\mathbf{H}}$ is a rescaled $\acute{\mathbf{H}}$, such that its $\sigma(\overline{\mathbf{H}})$ is confined to the interval [-1,1]. It is worth noting that Chebyshev polynomial terms $T_i(\overline{\mathbf{H}})$ are independent of ζ_j. Hence $\{\mathbf{y}(\zeta_j)\}_{j=1}^{l}$ can be generated in a single epoch.

Smith and co-workers [3, 9, 10] tridiagonalises $\acute{\mathbf{H}}$ first. The resulting Lanczos vectors $\{\mathbf{v}_i\}_{i=0}^{l-1}$ are then used to estimate the state filtered at ζ_j,

$$\mathbf{y}(\zeta_j) = \sum_{i=0}^{l-1} a_i(\zeta_j)\mathbf{v}_i, \tag{5}$$

In order to solve for the coefficients $a_i(\zeta_j)$, usually a linear problem (similar to that in SaiL, where the operator matrix is replaced with a shifted tridiagonal matrix) is solved. For more details on this and related schemes see Refs. [3, 11] and references there in.

Besides the Green operator, all the groups cited above employed other function operators, such as Dirac and Gaussian.

In realistic Hamiltonian systems, the resultant representative matrices are very large. This makes the memory requirement of $\{\mathbf{y}(\zeta_j)\}_{j=1}^{l}$ significant when l is of the order of 100's or more. Wall and Neuhauser cleverly circumvented this potential computer memory bottle-neck by building $\mathbf{Q}$ on the fly [12]. Once again, their algorithm was in the time-domain. Mandelshtam and Taylor [13, 14], Zhang and Smith[16], Chen and Guo [15], and Alacid *et al* [17] have developed equivalent time-independent versions. However, this modified FD, known as low-storage FD (LSFD), yields only

the eigenvalues of $\acute{\mathbf{H}}$. Eigenvector calculations demand knowledge of the set $\{\mathbf{y}(\zeta_j)\}_{j=1}^{l}$ whose regeneration can be computationally very expensive.

Skokov *et al* [44] computed all the vibrational energy levels of HOCl employing Cullum and Willoughby's Lanczos [37] and LSFD methods. They reported that the latter required about 4.8 times more CPU-time than the former. According to Skokov *et al*, this difference could be attributed to the implementation of a Fourier transform in the Lanczos code employed in their calculation, which was not a specific feature of the method. However, Huang and Carrington (HC) compared the performance of LSFD with a simple Cullum and Willoughby's Lanczos version for calculating the vibrational energy levels of a one-dimensional Morse oscillator and of the water molecule [45]. They cautiously concluded that a simple Lanczos method may be better suited than LSFD to calculating eigenvalues. The work of Zhang and Smith [46] seems to support this claim. Comparing the performances of LSFD and their LHFD (Lanczos homogenous filter diagonalisation) on computing the quasi-bound energies of HO_2, Zhang and Smith reported that LHFD required 2 to 6 fold fewer iterations (matrix-vector multiplications) than LSFD. In our view, however, it is not that clear that the LHFD approach improves on the LSFD algorithms or on the fast and cheap conventional triadiagonal matrix eigensolvers.

Overall, these studies suggest that a simple and easy to use Lanczos algorithm, a conventional eigensolver, is faster than LSFD.

2.3. FSM: Filtered or Lancsoz vectors as $\{\mathbf{s}_i\}_{i=1}^{l}$

The third and final set $\{\mathbf{s}_i\}_{i=1}^{l}$ for $\mathcal{S}$ is obtained by replacing the magnifier in $\{((\acute{\mathbf{H}}-\zeta\mathbf{I})^{-1})^i\mathbf{b}\}_{i=0}^{l-1}$ with a folder (*i.e.*, $\{((\acute{\mathbf{H}}-\zeta\mathbf{I})^{2})^i\mathbf{b}\}_{i=0}^{l-1}$ and at the same time imposing a condition that $((\acute{\mathbf{H}}-\zeta\mathbf{I})^{2})^i\mathbf{b}$ should skew $\mathbf{b}$ towards $\mathbf{u}_s$ of $(\acute{\mathbf{H}}-\zeta\mathbf{I})^2$. Alternatively a set of filtered states $\{\mathbf{y}(\zeta_j)\}_{j=1}^{l}$ is generated, where each $\mathbf{y}(\zeta_j)$ is close to $\mathbf{u}_s$ of $(\acute{\mathbf{H}}-\zeta_j\mathbf{I})^2$. In both cases and like SaiL and FD, orthogonalising the constructed vectors produces a good basis set to represent $\acute{\mathbf{H}}$ and yield $\mathbf{Q}$.

When the energy levels of $\acute{\mathbf{H}}$ are not dense, there is efficient software [51, 52] which can be used with SFM to form $\mathbf{Q}$ quite easily. In this case, the FSM approach becomes competitive with the other two methods. However, in general the inner part of the $\acute{\mathbf{H}}$ spectrum is dense rendering $(\acute{\mathbf{H}}-\zeta\mathbf{I})^2$ badly conditioned [26] so to avoid singularity good pre-conditioners are essential for its efficacy.

The FSM method was introduced by Wang and Zunger [25], although their original form of the method differs slightly from the descriptions given above. Wang and Zungerg developed FSM to find one eigenvalue at time — the lowest eigenvalue of $((\acute{\mathbf{H}}-\zeta\mathbf{I})^2$ which is equivalent to the eigenvalue of $\acute{\mathbf{H}}$ closest to ζ. For this they solved the problem via minimization of a functional.

In the discussions above it was assumed implicitly that the set $\{\mathbf{s}_i\}_{i=1}^{l}$ was computed accurately. This can be computationally expensive so the set is usually approximated. One consequence is that not all of the eigenvectors of $\mathbf{Q}$ converge to eigenpairs in the

vicinities of ζ. Henceforth l_c denotes the number of eigenvectors $\mathbf{Q}$ which also belong to $\acute{\mathbf{H}}$.

In summary, conceptually the scanning methods make use of the following self-evident idea: Any subspace $\mathcal{S}$ of $\mathcal{H}$, in which $\acute{\mathbf{H}}$ is defined, is associated with a portion of the energy spectrum. The methods, therefore, differ only in how to "project" out some components of $\mathbf{b}$ into $\mathcal{S}$ which entails: (1) The $f(\acute{\mathbf{H}}|\zeta_j)$ to employ (2) The strategy for performing $f(\acute{\mathbf{H}}|\zeta)\mathbf{b}$.

Both the FSM and SaiL methods require appropriate pre-conditioners [2, 21, 26], for which there is no existing systematic method for large Hamiltonian systems. The drawback of FD/LSFD is speed. Yu and Nyman [3] attributed this slowness to $f(\acute{\mathbf{H}}|\zeta_j)$ being expressed in polynomials in $\acute{\mathbf{H}}$: Since the polynomials are built up directly from $\acute{\mathbf{H}}$, the convergence rate of the spectrum computed by FD is strongly influenced by the convergence rate of the $\acute{\mathbf{H}}$ spectrum. Similar assertions were made by HC [21]. These arguments are important because the convergence rate of the Hamiltonian matrix $\acute{\mathbf{H}}$ spectrum is generally bad in the interior of the spectrum, the energy region for which FD was primarily developed [3]. However, we are aware of neither heuristic nor rigorous mathematical proofs tying the slowness of FD to the relative separation of the Hamiltonian matrix eigenvalues. In any case, it can be hard to avoid missing energy levels with FD [21, 47, 48], which often necessities repeated re-estimation of the filtered states.

In principle, all these shortcomings can be handled by inverting $(\acute{\mathbf{H}}-\mathbf{I}\zeta)^{-2}$, although in practice this has proved impossible [20]. Our proposed algorithm addresses this problem.

3. The CoBaE Algorithm

The proposed method is based on the basic concept of that $(\acute{\mathbf{H}}-\mathbf{I}\zeta)^{-2}$ can be obtained from $(\acute{\mathbf{H}}-\mathbf{I}\zeta)$. The derivation is given below but for simplicity of exposition we put most of the technical material in appendices.

3.1. Derivation

Theorem: *Any* $\acute{\mathbf{H}}$ *can be associated with a scaled real symmetric positive definite covariance matrix* of $(\acute{\mathbf{H}}-\mathbf{I}\zeta)$,*whose eigenvectors and those of* $\acute{\mathbf{H}}$ *are the same.*

Proof :—

Consider Eq. (2) as a linear fitting problem. If ζ is very close to an eigenvalue of the eigen-problem, for $\mathbf{w}$ to become the corresponding 'eigenvector', the following equation should be true

$$(\acute{\mathbf{H}}-\mathbf{I}\zeta)\mathbf{w}=\tau, \tag{6}$$

where $\tau \in \mathcal{R}^{n\times 1}(\neq \mathbf{0})$ is a residual vector. In essence $(\acute{\mathbf{H}}-\mathbf{I}\zeta)$ can be seen as the input data and τ as the outputs, where $\mathbf{w}$ contains the fitting parameters which are to be estimated. The *optimal value* of $\mathbf{w}$ can be found by minimising the cost/loss function

of Eq. (6), $\mathcal{J}(\mathbf{w}, \mathbf{H}-\mathbf{I}\zeta) = \Big[\tau - (\mathbf{H}-\mathbf{I}\zeta)\mathbf{w}\Big]^T\Big[\tau - (\mathbf{H}-\mathbf{I}\zeta)\mathbf{w}\Big]$, with respect to $\mathbf{w}$, *i.e.*, solving $\dfrac{\partial\mathcal{J}(\mathbf{w})}{\partial\mathbf{w}} = 0$ for $\mathbf{w}$. This yields the well known closed form solution

$$\Big[(\acute{\mathbf{H}}-\mathbf{I}\zeta)^T(\acute{\mathbf{H}}-\mathbf{I}\zeta)\Big]\mathbf{w} = (\acute{\mathbf{H}}-\mathbf{I}\zeta)^T\tau, \tag{7}$$

where $\frac{1}{n}\Big[(\acute{\mathbf{H}}-\mathbf{I}\zeta)^T(\acute{\mathbf{H}}-\mathbf{I}\zeta)\Big]$ and $\frac{1}{n}(\acute{\mathbf{H}}-\mathbf{I}\zeta)^T\tau$ are the covariance matrix of the input data, and the cross correlation between the input data and the output respectively [49, 50]. Therefore $\mathbf{B} = (\acute{\mathbf{H}}-\mathbf{I}\zeta)^T(\acute{\mathbf{H}}-\mathbf{I}\zeta)$ can be seen as *a scaled covariance matrix of the input data* (the shifted Hamiltonian matrix). Since $(\acute{\mathbf{H}}-\mathbf{I}\zeta)$ is symmetric and does not have zero eigenvalues, $\mathbf{B}$ is symmetric positive definite ($\sigma_j(\mathbf{B}) > 0$, $j = 1, 2, ..., n$), diagonalizable and invertible as well. For completeness these properties are proved in Appendix B.

Since $\mathbf{B}$ is diagonalizable, it can written as

$$\mathbf{B} = (\acute{\mathbf{H}}-\mathbf{I}\zeta)^2 = \mathbf{Z}\mathbf{D}\mathbf{Z}^T \tag{8}$$

where $\mathbf{D} \in \mathcal{R}^{n\times n} = diag(d_1 < d_2 <, ..., < d_n)$; $\mathbf{Z}^T\mathbf{Z} = \mathbf{I}$; the columns of $\mathbf{Z} \in \mathcal{R}^{n\times n}$ consist of the orthonormal eigenvectors $\mathbf{z}_j$ of $\mathbf{B}$; d_j are the corresponding eigenvalues [49, 50].

Eq. (8) can be rearranged to

$$\begin{aligned} \mathbf{D} \quad &= \mathbf{Z}^T(\acute{\mathbf{H}}-\mathbf{I}\zeta)^2\mathbf{Z} \\ \Rightarrow \mid \mathbf{D}^{\frac{1}{2}} \mid \; &= \mathbf{Z}^T\acute{\mathbf{H}}\mathbf{Z} - \mathbf{I}\zeta \\ \Rightarrow \sigma(\acute{\mathbf{H}}) \; &= \mathbf{Z}^T\acute{\mathbf{H}}\mathbf{Z} = \mid \mathbf{D}^{\frac{1}{2}} \mid + \mathbf{I}\zeta \end{aligned} \tag{9}$$

i.e., the columns of $\mathbf{Z}$ are also eigenvectors of $\acute{\mathbf{H}}$ as required.

The last line of Eq. (9) says that $\sigma_j(\acute{\mathbf{H}}) = \mid d_j^{\frac{1}{2}} \mid + \zeta$, that is finding the eigenvalues of $\acute{\mathbf{H}}$ in the vicinity of ζ ($eHv\zeta$) means yielding the smallest eigenpairs of $\mathbf{B}$. However, a close inspection of the proof reveals that it is actually similar to the FSM method whose major difficulty was mentioned earlier. Fortunately $\mathbf{B}$ is invertible hence, in principle, $\mathbf{B}^{-1} = \mathbf{Z}\mathbf{D}^{-1}\mathbf{Z}^T$ which simplifies to

$$\begin{aligned} \mid \mathbf{D}^{-\frac{1}{2}} \mid \; &= (\sigma(\acute{\mathbf{H}}) - \mathbf{I}\zeta)^{-1} \\ \Rightarrow \sigma(\acute{\mathbf{H}}) = \; &\mid \mathbf{D}^{\frac{1}{2}} \mid + \mathbf{I}\zeta \end{aligned} \tag{10}$$

Eq. (10) shows that the larger $\mid d_j^{-\frac{1}{2}} \mid$ is, the closer $\sigma_j(\acute{\mathbf{H}}) = \mid d_i^{\frac{1}{2}} \mid + \zeta$ becomes to ζ. Thus, as $\mid d_1^{-\frac{1}{2}} \mid > \mid d_2^{-\frac{1}{2}} \mid > .... > \mid d_n^{-\frac{1}{2}} \mid$, the task of obtaining $eHv\zeta$ amounts to calculating the few eigenvectors of $\mathbf{B}^{-1}$ with the largest eigenvalues. Recall the smallest eigenvalue of $(\acute{\mathbf{H}}-\mathbf{I}\zeta)^2$ and the largest eigenvalue of $(\acute{\mathbf{H}}-\mathbf{I}\zeta)^{-2}$ have the same eigenvector [30]. Eq. (10) also implies that the largest eigenvalues of $\mathbf{B}^{-1}$ are well separated when $\mid d_1^{\frac{1}{2}} \mid$, $\mid d_2^{\frac{1}{2}} \mid$,...,and $\mid d_l^{\frac{1}{2}} \mid$ of $\mathbf{B}$ near ζ are quite close to each other. In other words if $\mathbf{B}^{-1}$ is known, generating $\{\mathbf{s}_i\}_{i=1}^{l}$ is equivalent to approximating the few largest eigenvectors of $\mathbf{B}^{-1}$.

Here we propose a scanning algorithm based on estimating the few dominant eigenpairs of $\xi^2\mathbf{B}^{-1}$ without inverting $\mathbf{B}$ directly. The scheme consists of three main steps: (1) The Inversion Step (IS), calculating $\xi^2\mathbf{B}^{-1}$, (2) Subspace Generating Step (SGS), constructing $\{\mathbf{s}_i\}_{i=1}^{l}$, and (3) Construction and diagonalization Step (CDS), obtaining and diagonalising $\mathbf{Q}$.

Steps 1 and 2 could be amalgamated, but for clarity we treat them separately. The significance of ξ^2 is discussed in Section 4. In the meantime, we note that $\xi^2 \in \mathcal{R}^+$, a positive real number.

3.2. Inversion and Subspace Generation Steps

As discussed in Appendix A, Eq. (A.11), $\xi^2\mathbf{B}^{-1}$ can be expressed as

$$\xi^2\mathbf{B}^{-1} = \mathbf{I} - \sum_{j=1}^{n} \frac{\mathbf{g}_j\mathbf{g}_j^T}{\varsigma_j} \tag{11}$$

Given $\xi^2\mathbf{B}^{-1}$, the set $\{\mathbf{s}_i\}_{i=1}^{l}$ can be obtained with the basic Krylov subspace algorithms that we discussed in the previous section, or with any of its sophisticated variants, such as the basic Lanczos schemes, Implicitly Restarted Lanczos Methods [51], Simultaneous Iterations algorithms [52], to name but a few. Without loss of generality, as an illustration, a basic Krylov algorithm is employed for generating the small subspace where, in the vicinity of ζ, the eigenpairs pertaining to $\acute{\mathbf{H}}$ reside.

The Krylov sequence $\{(\xi^2\mathbf{B}^{-1})^m\mathbf{b}\}_{m=0}^{l-1}$ is formed via the following two term recurrence

$$\nu_{m+1} = \nu_m - \sum_{j=1}^{n} \frac{\mathbf{g}_j\mathbf{g}_j^T}{\varsigma_j}\nu_m, \;\; m = 0, 1, 2, ..., l-1 \tag{12}$$

where $\nu_m \in \mathcal{R}^{n\times 1}$, $\nu_0 = \mathbf{b}$; $\mathbf{g}_j$ and ς_j are as defined in Eq. (A.15).

$\acute{\mathbf{H}}$ is then expressed in the orthogonalised version of $\{\nu_m\}_{m=0}^{l-1}$, $\{\mathbf{s}_m\}_{m=1}^{l}$, to obtain $\mathbf{Q} = (\mathbf{s}_m, \acute{\mathbf{H}}\mathbf{s}_{m'})_{m,m'=1}^{l}$ whose diagonalisation yields the l_c eigenpairs of $\acute{\mathbf{H}}$ in the neighbourhood of ζ.

This completes the derivation of the proposed method which we call Covariance Based Eigensolver (CoBaE).

4. Implications

Drawing upon the above findings, some possible implications are discussed below.

4.1. Computational: Cost, the role of ξ^2, and CoBaE stability

To determine the computational cost of the CoBaE algorithm we use η to denotes the average number of non-zero elements per $\mathbf{g}_j$, λ is defined as the lower bound for a given value of ζ^2, see Appendix A, and l gives the number of matrix-vector multiplications required to construct the basis set spanning the subspace.

The CPU-time cost of the inversion step (IS), $\mathbf{g}_j = \mathbf{x}_j - \sum_{i=1}^{j-1} \frac{\beta_{ij}}{\varsigma_i}\mathbf{g}_i$ and $\mathbf{g}_j = \mathbf{x}_j - \sum_{i=j-\lambda}^{j} \frac{\beta_{ij}}{\varsigma_i}\mathbf{g}_i$, are given by $\frac{(\lambda-1)}{2}[2\eta+(\lambda-1)\eta]$ and $\approx O[(\eta-1)(\lambda\times\eta)]$ respectively; $j\times\eta$ and $\lambda\times\eta$ words respectively are their memory requirements. Storing $\{\mathbf{g}_j\}_{j=1}^{n}$ costs about $n\times\eta$ words.

In the subspace generation step (SGS), the main CPU-time cost is $\approx O(l\times n\times\eta)$ and is due to performing the matrix-vector multiplication $\nu_{m+1} = \nu_m - \sum_{j=1}^{n} \frac{\mathbf{g}_j\mathbf{g}_j^T}{\varsigma_j}\nu_m,\ m = 0, 1, 2, ..., l-1$. The memory requirements depend mainly on the subspace generating algorithm (SGA) employed.

The computational cost of constructing and diagonalising $\mathbf{Q}$ step (CDS) is negligible when the generated basis set is not orthogonalised, but instead the Singular Value Decomposition of the basis set overlap matrix is computed.

For CoBaE to be efficient computationally, l, λ and in particular η must be much less than n.

Substituting Eq. (A.15) into Eq. (11) results in

$$\xi^2\mathbf{B}^{-2} = \mathbf{I} - \sum_{j=1}^{n}\frac{\mathbf{x}_j\mathbf{x}_j^T}{\varsigma_j} + \sum_{j=1}^{n}\sum_{i=1}^{j-1}\frac{\beta_{ij}}{\varsigma_j\varsigma_i}(\mathbf{g}_i\mathbf{x}_j^T + \mathbf{x}_j\mathbf{g}_i^T) - \sum_{j=1}^{n}\sum_{ik=1}^{j-1}\frac{\beta_{ij}\beta_{kj}}{\varsigma_j\varsigma_i\varsigma_k}\mathbf{g}_j\mathbf{g}_k^T$$

where $\varsigma_{j,k,i} = \xi^2 + \mathbf{x}_{j,k,i}^T\mathbf{g}_{j,k,i}$.

In the case $\xi^2 + \mathbf{x}_{j,k,i}^T\mathbf{g}_{j,k,i} \approx \xi^2$, the above equation simplifies in matrix form to

$$\xi^2\mathbf{B}^{-1} = \mathbf{I} - \Big[\frac{(\acute{\mathbf{H}} - \zeta\mathbf{I})^2}{\xi^2} - \frac{\Upsilon\Upsilon^T}{(\xi^2)^2} + \frac{\Xi\Xi^T}{(\xi^2)^3}\Big]$$

$\Rightarrow$

$$\xi^2\mathbf{B}^{-1} \approx \mathbf{I} - \frac{(\acute{\mathbf{H}} - \zeta\mathbf{I})^2}{\xi^2} = \mathbf{I} - \xi^{-2}\acute{\mathbf{H}}^2 + 2\xi^{-2}\zeta\acute{\mathbf{H}} - \xi^{-2}\zeta^2\mathbf{I} \tag{13}$$

A closer look at Eq. (13) reveals that the largest eigenvalues of $\xi^2\mathbf{B}^{-1}$, those pertaining to its dominant eigenvectors, are close to 1.0 (or to ϵ, see Eq. (20) below.) In subspace generating approaches, it is a basic knowledge that converging tightly clustered eigenvalues takes a significant amount of matrix-vector multiplications. In other words increasing ξ^2 can result in large number of matrix-vector multiplications, *i.e.*, a large value of l. On the other hand by induction, reducing the value of ξ^2 causes l to decrease.

However, the reduction of l, *i.e.*, ξ^2, is achieved at the cost of increased λ to maintain the positive definiteness of $\xi^2\mathbf{B}^{-1}$, see Appendix A. A disadvantage of the rise in the λ value is a consequent increase in CPU-time requirements of IS which is quadratically or linearly dependent on λ. Nevertheless as λ has an upper limit of $\leq n-1$, the worse-case time requirement of IS is always less than $O(n^2)$. One may therefore not worry about the value of λ becoming large. Furthermore, as mentioned in Appendix A, due to the linearity of the Hamiltonian systems, even if the covariance matrix loses its positive definiteness, one can still extract the desired eigenpairs of the Hamiltonian matrix from $\xi^2\mathbf{B}^{-1}$ spectrum, provided that all the dominant eigenvalues of the covariance matrix remain positive.

Another consequence of small ξ^2 is that it indirectly increases the value of η, see above and Eq. (A.16). CoBaE scales unfavourably with η, in particular its memory requirements. Fortunately, the Hamiltonian matrices that we are mainly concerned with are sparse or diagonally dominant. Thus, in principle, the value of η can be controlled with an appropriately chosen ε as described in Appendix A.

To summarise it, both in principle and in practice, small values of l, η, and, to certain extent, λ are attainable rendering the proposed method computationally efficient.

Now we comment on the statement that in SGS the memory requirements depend mainly on the subspace generating algorithm (SGA) one employs. In Krylov algorithms, it is necessary to store the generated sequence $\{\nu_{m+1}\}_{m=0}^{l-1}$. In IRLM or Simultaneous Iterations methods, the size of the basis set is pre-defined, and therefore memory requirements are not directly determined by l. When a Lanczos algorithm is employed as the SGA, only the storage of three vectors are required. Ultimately the values of l and n determine which method should be used as the SGA.

Note that in the case where the basic Lanczos method is the SGA, CoBaE yields only the eigenvalues of the Hamiltonian matrix in the vicinities of ζ. However, as we have just described, for a small ξ^2, the order of propagation (l) is small. Thus, unlike LSFD and LHFD, the computation of the corresponding eigenvectors can be relatively much cheaper.

Finally we briefly discuss the numerical stability of the method. In Appendex A, we noted that the Sherman-Morrison scheme may become numerically unstable owing to accumalative residual errors. As detailed in [53, 54], this can be mitigated in a number of different ways by: Adding artificial noises to Eq. (A.6), computing the square root of $\xi^2\mathbf{B}^{-1}$ instead of $\xi^2\mathbf{B}^{-1}$, or using a higher precision floating point arithmetic; the latter two options are quite easy to implement and computationally efficient. In the test results presented in this article double precision floating point arithmetic in Fortran was used.

The above semi-qualitative analysis of the computational role of ξ^2 in CoBaE was tested and some of the results are given in Section 5. A full computational analysis including parallelization of the algorithm, which is highly parallelisable, will be given elsewhere [65].

4.2. Physical interpretation

One of the important consequence is that CoBaE can extract not only eigenpairs, but $\acute{\mathbf{H}}$ itself from quantum states which are prepared experimentally. Another possible implication is the apparent connections between the scanning algorithms and quantum statistics, and in particular with quantum control.

4.2.1. Mixed Quantum State: In the absence of other information, a Gaussian distribution with a unit variance $\mathbf{I}(\in \mathcal{R}^{n\times n})$ was implicitly invoked to describe possible variations in τ resulting in the least square problem given in Eq. (6) [50, 61]. In this

section we explore the variations in τ and subsequent variations in the 'eigenvector' in Eq. (6) a bit more carefully to reveal much more of the structure underlying the scanning schemes and any physical insight this might offer.

If τ is considered as a n-dimensional random vector which has a multivariate normal distribution density function with a covariance matrix $O \in \mathcal{R}^{n\times n}$, the linearity of $(\acute{\mathbf{H}} - \mathbf{I}\zeta)$ ascertains that the parameter vector in Eq. (6) should also be characterised by a multivariate Gaussian probability density function with a covariance matrix $\mathbf{F} \in \mathcal{R}^{n\times n}$ [60]. The relationship between the two covariance matrices is given by the law of covariances [50, 60] as

$$\mathbf{F} = ((\acute{\mathbf{H}} - \mathbf{I}\zeta)^T)^{-1}\mathbf{O}(\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1} \tag{14}$$

which means $\mathbf{O} \rightarrow 0(\infty) \Rightarrow \mathbf{F} \rightarrow 0(\infty)$.

Eq. (14) simplifies to $\mathbf{F} = \xi_i^2(\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-2} = \xi_i^2\mathbf{B}^{-1}$ if $\mathbf{O} = \xi_i^2\mathbf{I}$ $(i = 1, 2, ..., n)$, *i.e*, there is no correlation between the errors in the different components of τ. ξ_i^2 denotes the variance in the i^{th} element of τ. $\mathbf{F} = \xi_i^2\mathbf{B}^{-1}$ further reduces to $\mathbf{F} = \xi^2\mathbf{B}^{-1}$ [cf. Eq. (11)] when all the variances are equal $\xi_i^2 = \xi^2$, $\forall i$.

Evidently the unknown quantum state, the 'eigenvector' associated with ζ in Eq. (6) is not a point (pure state) in the Hilbert space, but a distribution, mixed state, of an ensemble of pure states whose classical probability density function is Gaussian with a covariance matrix $\mathbf{F}$ where $\mathbf{F}$ is proportional to the square of the Green operator, $\xi^2(\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-2}$. Moreover, the optimal value of the random vector $\mathbf{w}$ in Eq. (6) can be seen as an instance and corresponds to a pure state in the Hilbert space of $(\acute{\mathbf{H}} - \mathbf{I}\zeta)$.

From mathematical perspective a Gaussian probability density function (pdf) is completely specified by its first and second moments. The latter not only quantifies the uncertainties/variances, but it also determines the shape of the pdf in state space. Thus $\xi^2\mathbf{B}^{-1}$ decides how isotropic/anisotropic the pdf, which characterises the mixed state, is in Hilbert space. Isotropicity suggests the mixed state is equally likely to be found anywhere in Hilbert space (total ignorance) — that is, the shape of the pdf in Hilbert space is that of a hypersphere whose all principal axes (the eigenvectors of $\xi^2\mathbf{B}^{-1}$) are equal. Conversely a pdf with a small $\xi^2\mathbf{B}^{-1}$ is sharply anisotropic, a hyperellipsoid.

In other words the mixed state associated with the shifted Hamiltonian operator is highly likely to be found in a small region/subspace of Hilbert space. Besides the notion of strong anisotropicity indicates the dominance of few principle axes (eigenvectors) of $\xi^2\mathbf{B}^{-1}$, and in fact this small set of eigenvectors approximately span the subspace [62].

Obviously the statistics (in particular the covariance) of the unknown mixed quantum state can be extracted from $\acute{\mathbf{H}}$, ζ and $\mathbf{O}$. However, knowledge of $\acute{\mathbf{H}}$, ζ and $\mathbf{O}$ is not always necessary. For instance, the covariance matrix of the mixed state can be estimated from experimental data. Furthermore, we have shown that given $\acute{\mathbf{H}}$, ζ and $\mathbf{O}$, CoBaE obtains the eigenpairs of $\acute{\mathbf{H}}$ in the vicinities of ζ from $\xi_i^2\mathbf{B}^{-1}$, the covariance matrix of the mixed state.

Therefore, in principle, CoBaE can extract not only its eigenpairs, but $\acute{\mathbf{H}}$ itself from a covariance matrix associated with an ensemble of pure states which might be prepared

experimentally.

4.2.2. Potential Energy Surface Probe Since the covariance matrix is a squared quantity $\xi^2(\acute{\mathbf{H}} - \zeta\mathbf{I})^{-2}$, $d_j^{\frac{1}{2}}$ can carry either a positive or negative sign. Thus Eq. (10) produces a sign pattern signature (SPS) unique to the given $\acute{\mathbf{H}}$ spectrum. For a simple vibrator this pattern is determined by the potential energy surface (PES). For instance the SPS and $\mid d_j^{\frac{1}{2}} \mid$ for the spectrum in an energy region where the PES is harmonic is different than those for a PES which is Morse in nature — we have found that the SPS of the Harmonic and Morse spectra are given by $(-1)^{j+1}$ and $(-1)^j$ respectively, where $j = 1, 2, ..., l$. This means that our scanning method could be used as a probe of the underlying PES.

4.3. Encapsulation of FD, SaiL and FSM

Encapsulation means that although superficially it may appear that the CoBaE and FSM, SaiL and FD methods give estimates for different states for a given ζ, a closer look reveals the states are in fact the same. It is a mixed state which is likely to be found in a subspace spanned by a small set of eigenvectors — effectively it is this set of vectors the four methods are estimating. The eigenvectors are the principal axes of the covariance matrix associated with the state when ξ^2 equals unity.

Rearranging Eq. (7) gives the optimal value of the state as $\mathbf{w} = \mathbf{B}^{-1}\mathbf{t}$, where $\mathbf{t} = (\acute{\mathbf{H}} - \zeta\mathbf{I})\tau$. In a spectral representation, this state is rewritten as

$$\mathbf{w} = \mathbf{Z}\mathbf{D}^{-1}\mathbf{Z}^T\mathbf{t} \tag{15}$$

Although $\mathbf{B}^{-1}$ is n-dimensional matrix, its l_c eigenvectors $\{\mathbf{z}_i\}_{i=1}^{l_c}$in the subspace $\mathcal{S}$, which pertain to the l_c largest eigenvalues $\{d_i^{-1}\}_{i=1}^{l_c}$of the matrix are far more significant, in their information content, than the other eigenpairs. Thus one can approximate $\mathbf{B}^{-1}$ by [62],

$$\begin{aligned} \mathbf{B}^{-1} &\approx \sum_{i=1}^{l_c} d_i^{-1}\mathbf{z}_i\mathbf{z}_i^T \\ \Rightarrow \mathbf{w} &\approx \sum_{i=1}^{l_c} a_i\mathbf{z}_i \end{aligned} \tag{16}$$

where $a_i = d_i^{-1}(\mathbf{z}_i^T\mathbf{t})$.

Recall that while the matrices $(\acute{\mathbf{H}} - \mathbf{I}\zeta)^k$ ($k = 2, -2, 1, or\ -1$) have the same eigenvectors $\mathbf{z}_i$, their corresponding eigenvalues are different: d_i, d_i^{-1}, $d_i^{\frac{1}{2}}$, or $d_i^{-\frac{1}{2}}$ respectively [30, 31]. Thus considering each method in turn.

(i) *SaiL:* Solving Eq. (6)

$$(\acute{\mathbf{H}} - \mathbf{I}\zeta)\mathbf{w} = \tau \tag{17}$$

amounts to solving $(\mathbf{Z}\mathbf{D}^{\frac{1}{2}}\mathbf{Z}^T)\mathbf{w} = \tau \Rightarrow \sum_i^n d_{i=1}^{\frac{1}{2}}\mathbf{z}_i\mathbf{z}_i^T\mathbf{w} = \tau$, *i.e.*, $\mathbf{w} \approx \sum_{i=1}^{l_c} a_i\mathbf{z}_i$, where $a_i = d_i^{-\frac{1}{2}}(\mathbf{z}_i^T\tau)$. In this case the eigenvectors of $(\acute{\mathbf{H}} - \mathbf{I}\zeta)$ with the smallest eigenvalues contribute to the estimation of $\mathbf{w}$ the most.

(ii) *FD:* Here $\mathbf{w}$ is the filtered state $\mathbf{y}(\zeta_j)$ in Eq. (4). Thus

$$\mathbf{w} = (\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1}\tau = \sum_{i=1}^{K} q_i(\zeta_j)T_i(\bar{\mathbf{H}}|\zeta_j)\tau \tag{18}$$

is equivalent to $\mathbf{w} = \sum_i^n d_{i=1}^{-\frac{1}{2}}\mathbf{z}_i\mathbf{z}_i^T\tau \Rightarrow \mathbf{w} \approx \sum_{i=1}^{l_c} a_i\mathbf{z}_i$ where $a_i = d_i^{-\frac{1}{2}}(\mathbf{z}_i^T\tau)$. In this case the eigenvectors of $(\acute{\mathbf{H}} - \mathbf{I}\zeta_j)^{-1}$ with the largest eigenvalues (in magnitude) contribute the most to the estimation of $\mathbf{w}$. As explained in Section 1, these are the eigenvectors in $\mathbf{b}$ (here τ), which are magnified the most by $(\acute{\mathbf{H}} - \mathbf{I}\zeta_j)^{-1}$.

(iii) *FSM:* As $\mathbf{B} = \mathbf{Z}\mathbf{D}\mathbf{Z}^T$, its l_c smallest eigenvectors contribute most to the estimation of $\mathbf{w}$. As described in Section 2, these are eigenvectors that the FSM methods incidentally computes.

Below we present test results for the method, but before that we address the initializations of $\mathbf{P}_0$ in Eq. (A.8) which we skipped at the time.

Unfortunately the appropriate *a priori* information about the quantum state is not generally available. Thus in CoBaE $\mathbf{P}_0$ is initialised with the matrix $\epsilon\mathbf{I}$, where ϵ is set arbitrarily to a large positive real number: Positive because it denotes a variance, and large because where in the Hilbert space to look at for the quantum state is unknown *a priori*. Replacing $\mathbf{I}$ with $\epsilon\mathbf{I}$ in Eqs. (11) and (A.15) results in the general form of the proposed algorithm.

$$\xi^2\mathbf{B}^{-1} = \epsilon\mathbf{I} - \sum_{j=1}^{n} \frac{\epsilon^2\mathbf{g}_j\mathbf{g}_j^T}{\varsigma_j} \tag{19}$$

$$\mathbf{g}_j = \mathbf{x}_j - \begin{cases} \sum_{i=1}^{j-1} \frac{\beta_{ij}}{\varsigma_i}\epsilon\mathbf{g}_i & \text{for } j-1 \le \lambda \le n-1 \\ \sum_{i=j-\lambda}^{j} \frac{\beta_{ij}}{\varsigma_i}\epsilon\mathbf{g}_i & \text{for } \lambda < j \le n \end{cases} \tag{20}$$

whereas $\varsigma_j = \xi^2 + \mathbf{x}_i\epsilon\mathbf{g}_i$.

Once again, the uncertainties in τ are also generally unknown *a priori*. Hence ξ^2 is treated as a parameter in CoBaE whenever ξ^2, or $\mathbf{O}$, is not available.

5. Analyses and Test Results

As a sample application of the method we use the vibrational energy levels of the HOCl molecule to demonstrate CoBaE. The Hamiltonian operator is real symmetric and defined in Jacobi co-ordinate (R, r, θ) system [4]. In this work the co-ordinates were chosen such that R is the distance of Cl from the centre of mass of OH, r is the OH

inter- nuclear distance and θ is the angle between R and r. The Hamiltonian operator was represented in Discrete Variable Representation (DVR), in particular the sequential diagonalisation and truncation version [63, 64]. In this representation, motions in R and r are expressed in radial DVR functions (β and α respectively) based on Morse-oscillator-like functions which have variationally optimisable parameters (r_e, ω_e and D_e) [64, 66]. The bending vibration, motion in θ, is expressed in angular DVR functions (γ) based on (associated) Legendre polynomials.

In this work the primitive DVR basis were 96, 45 and 60 DVR functions β, α and γ in R, r, and θ, respectively. The variational parameters r_e, ω_e and D_e, were optimised for the two radial motions [66]. For r they were set to $4.8295a_0$, $0.01654E_h$, and $0.00230E_h$ respectively; for R to $8.980a_0$, $0.000080E_h$, and $0.0000554E_h$ [66]. The primitive DVR basis set and the variational parameters are those employed in Ref.[4] where all the energy levels were found to be stable and well converged with these parameters and basis.

After a single sequential diagonalisation and truncation step (as described in Ref.[4]), a dense 3D-Hamiltonian matrix $\mathbf{\acute{H}}$ with a dimension of 9600 was constructed. The denseness of the matrix, in particular, made this test case realistically challenging for CoBaE. This allowed us to verify the validity of some of our speculation on the roles of ξ^2 and ε in the method.

In the following analyses we concentrated on the most dense, and hence difficult, part of the vibrational spectrum, an energy window centred at $\zeta = 19267.907$ cm^{-1}. This was about 20 cm^{-1} below the dissociation threshold (D_0) of the potential energy surface (PES) employed, that of Skokov *et al* [67]. All the calculations were performed on a three year old Intel-Pentium 3.60 GHz PC and 1 GB RAM.

In both the paragraph above and the rest of the discussion, the quoted energy values are band origins.

A threshold ε was pre-selected to 0.0 or 1×10^{-5}. The covariance matrix $\mathbf{P}_0$ of Eq. (20) was initialised with $\epsilon = 100000.0$. The small l-dimensional subspaces were generated employing basic ARPACK, an algorithm based on the Implicitly Restarted Lanczos Methods (IRLM) [51]. It is arguably the most sophisticated method for generating subspaces, but at the cost of increased computer memory where storing two vectors, **vv**(n,ϱ) and **work**(ϱ^2+8ϱ), are its main core-memory overheads.

In order to put things into prospective, to converge the energy levels in the window above, in its basic form the ARPACK suite would require ϱ of 1700, but a much smaller ϱ value for calculating the eigenpairs at the lower end of the spectrum as the energy levels are well separated. Moreover, ϱ does not only determine the memory overhead, but it also affects the performance of ARPACK: The larger ϱ is the smaller the number of matrix-vector multiplication operations, l, required which leads to a reduction in CPU-time requirements. For full details of ARPACK, see Ref. [51].

Note that unlike the basic Krylov approach we used to illustrate CoBaE, in ARPACK it is necessary to pre-define the dimension of the subspace, ϱ. In this work ϱ was set to 50, not only to reduce memory overhead, but also to test how well IS of

CoBaE separates the eigenpairs of $\xi^2\mathbf{B}^{-1}$ near ζ. Finally the convergence tolerance of ARPACK was set to 1×10^{-5} and no effort was to made to optimise the seeding vector, it was set to one.

The new method was tested against FSM. To a lesser extent the method was also compared with the performance of FD and the Lanczos algorithm. Our results are summarised in Table 1. We started by setting ε to zero.

To test how varying ξ^2 affects the speed of CoBaE, ξ^2 was set to 0.0025 and 400.0 respectively. With $\xi^2 = 0.0025$, IS took 1.75 mins while SGS required 48.62 mins and 125 matrix-vector multiplication operations (MVO) to converging the 10 energy levels closest to 19267.907 cm^{-1} to within 0.01 cm^{-1} of the exact results or better. Here none of the terms in $\mathbf{g}_j = \mathbf{x}_j - \sum_{i=1}^{j-1} \frac{\beta_{ij}}{\varsigma_i}\mathbf{g}_i$ was dropped, *i.e*, instead of Eq. (20), Eq. (A.10) was employed. With $\xi^2 = 400.0$, only 10 (*i.e.*, λ=10) terms were required to construct $\mathbf{g}_j$ (see Eq. (20)). IS took 0.23 mins while SGS required 31878 MVO and 327.91 mins to converge the same number of energy levels to similar accuracy.

Columns 2 and 3 of Table 1 give the 10 vibrational energy levels in the immediate vicinity of 19267.907 cm^{-1} that resulted from the two calculations. Clearly these energies are in excellent agreement with the energy levels yielded by a conventional diagonaliser, the NAG subroutine F02ABF [68].

Keeping λ to 10, but setting ξ^2 to 0.0025, we repeated the last calculation. As expected, see the discussion following Eq. (A.14), the covariance matrix $\xi^2\mathbf{B}^{-1}$ lost its positive definiteness — that is, all the largest eigenvalues (in magnitude) of $\xi^2\mathbf{B}^{-1}$ became negative.

Next both ξ^2 and λ were changed to 1.0 (= 400/400) and 800 ($=4\sqrt{400} \times 10$) respectively. IS took 7.92 mins whereas SGS required 10.18 mins and 959 MVO. The fourth column of Table 1 shows that the same 10 energy levels resutling from this calculation are again in excellent agreement with the vibrational energy levels yielded by F02ABF.

These results appear to confirm our speculation that ξ^2 not only determines the λ values, but also influences the speed of the different components of CoBaE. However, so far the set $\{\mathbf{g}_j\}_{j=1}^n$, which was dense, was retained in computer memory in all the calculations.

Next the threshold ε was arbitrarily set to 1×10^{-5} and then the last calculation was repeated. Over 90% of the contents of the set $\{\mathbf{g}_j\}_{j=1}^n$ became zero. Moreover, the only notable differences between the two runs were in the accuracy of the energy levels as column 5 of Table 1 shows.

Given the fact that the Hamiltonian matrix was dense and the set $\{\mathbf{g}_j\}_{j=1}^n$ was made sparse arbitrarily, some deterioration in accuracy of the computed spectrum was understandably inevitable. However, it was quite surprising that the level of deterioration in the accuracy of the relevant 10 energy levels was actually far less than expected. The energy levels are within 0.5 cm^{-1} of the exact results. This observation was found to be true over the entire vibrational spectrum.

Again these results appear to verify some of the inferences made above. In particular

that $\mathbf{g}_j$ can be rendered sparse even when the Hamiltonian matrix is dense and (not necessarily) diagonally dominant while maintaining the accuracy of the Hamiltonian operator spectrum.

The lower energy windows of the system were far much less costly than the energy window described above.

FSM was employed to compute the 10 energy levels in the vicinities 19267.907cm^{-1}. Like CoBaE ARPACK was used as SGA, where both v and the convergence tolerance were the same as in CoBaE. FSM required 63528 MVO and 603.25 mins. The results are as given in column 6 of Table 1. To our knowledge this was the first time FSM was used to calculate vibrational energy levels of a molecule.

Finally the 10 eigenstates were computed using FD, as given in Eq.(4), where the the range of the energy window was set to [19224.328cm^{-1},19306.535cm^{-1}]. 26000 Chebyshev terms — that is, 26000 MVO — and a total of 50 filtered states were required. In this calculation, the singular values of the filtered state overlap matrix were computed with a LAPACK routine, **dgesvd**. All the significant eigenvectors (25 in total) of the overlap matrix were then used for the construction of $\mathbf{Q}$ whose diagonalisation yielded the required 10 energy levels, see column 7 of Table 1. The calculation took 319.40 mins.

Evidently the results obtained with CoBaE, FD and FSM are in excellent agreement. The computaional performances of FD and FSM (which are equal or about 2 times slower than CoBaE when ξ^2 was set to 400, the worst case scenario of CoBaE) are in line with the prediction of Eq. (13). In all other cases, both FD and FSM are far much slower than CoBaE. For instance, FD required respectively about 18 and 27 times more CPU time and MVO than the best case of the CoBaE examples given above.

In all the computations in which ε was 1×10^{-5}, we did not make any effort to drop the zero components of $\{\mathbf{g}_j\}_{j=1}^n$, *i.e*, over 90% of the set, from the calculations. This significantly overestimates the CPU-time requirements of CoBaE indeed. A more robust comparison will be given elsewhere.

6. Conclusion

The method proposed here is simple, versatile and generic. We expect that it will be applied for spectral scanning, be it with energy or otherwise. Furthermore the scheme gives a different perspective on scanning algorithms and physical insight into solving Hamiltonian systems. The test case we studied clearly supports some of the inferences made during the derivation. For example, the test shows that even when the Hamiltonian matrix is dense and cannot be made sparse, CoBaE can still work well and yield the spectrum of the matrix with acceptable accuracy.

Here we have focused on the derivation of the algorithm, applications of the method will be given elsewhere [65], where a rigorous comparison against other scanning algorithms, in particular LSFD, will be given in detail. However, we point out two observations that in our view warrant further research.

The results from Eq. (13), CoBaE and FSM with ξ^2 of 400.0 and 1.0 respectively, appear to confirm Yu and Nyman (YN)'s analysis concerning the slowness of the FD scheme [3, 21], see Section 2 of this article. Although $\xi^2\mathbf{B}^{-1}$ is constructed (or rather approximated) directly from $\acute{\mathbf{H}}$, the cause of the slowness which the equation describes is different from what YN attributed to the low convergence rate of FD. They argue that the speed of FD depends on the spectral range and the separation of the desired eigenvalues of the Hamiltonian matrix since FD's function operators are directly built from the Hamiltonian matrix.

However, in the case of CoBaE, the convergence rate of $\xi^2\mathbf{B}^{-1}$ spectrum, as given in Eq. (13), is independent of the spectral range of the Hamiltonian matrix. Instead it depends on the relative separtion of the dominant eigenvalues of $\xi^2\mathbf{B}^{-1}$. (Recall that a large ξ^2 tightly clusters these eigevalues around ϵ.) Thus, here the clustering effect of ξ^2, rather than the covariance operator being built directly from polynomials in $\acute{\mathbf{H}}$, is the source of the slowness. Even though the eigenpairs of interest are tightly clustered, nevertheless they are significantly dilated and well separated from the rest of $\xi^2\mathbf{B}^{-1}$ spectrum. By making use of these facts one might be able to develop mathematical techniques capable of reducing the number of matrix-vector multiplications and speed up CoBaE in the scenario where ξ^2 is noticeable large.

The second point is that scanning algorithms are similar in spirit to, and might be relevant for, Quantum Control problems [29], where one controls the dynamics or measurements of quantum systems via the manipulation of external parameters. When treated as parameters, ξ^2 and ζ control (in a pseudo sense) the Hamiltonian and in doing so steer the physical system from a probabilistically known initial state to a target state with certain probability. Further studies might shed some light on this.

Acknowledgments

HYM and RCG thank Unilever for funding.

Appendix A. Inversion Step

In Section 2 we show that

$$\mathbf{B} = (\acute{\mathbf{H}} - \mathbf{I}\zeta)^T(\acute{\mathbf{H}} - \mathbf{I}\zeta) \tag{A.1}$$

This allows us to express $\mathbf{B}$ in terms of the contents of the shifted Hamiltonian matrix

$$\begin{aligned}\mathbf{B} &= \sum_{i=1}^{n} \mathbf{x}_i\mathbf{x}_i^T \\ &= \sum_{i=1}^{n-1} \mathbf{x}_i\mathbf{x}_i^T + \mathbf{x}_n\mathbf{x}_n^T\end{aligned} \tag{A.2}$$

where $\mathbf{x}_i \in \mathcal{R}^{n\times 1}$ is the i^{th} column of $(\acute{\mathbf{H}} - \mathbf{I}\zeta)^T$.

Table 1. Results for the 10 vibrational energy levels near the dissociation limit of HOCL. The energy levels are band origins in cm^{-1}. "SN" means state number; "Exact" stands for results from a conventional, full matrix eigensolver [68]; other results are given as a difference from the exact values. The matrix has $n = 9600$; ε denotes the threshold value; *NA* indicates that Eq. (A.10), instead of Eq. (20), was used to generate $\mathbf{g}_j$. λ, FSM and FD are as described in the text.

	ε= 0.0			ε= 10^{-5}	FSM	FD	Exact
	ξ^2=400.0	ξ^2=0.0025	ξ^2=1.0	ξ^2=1.0	ξ^2=1.0		
SN	λ=10	λ=NA	λ=800	λ=800			
809	0.000	0.000	0.009	0.109	0.000	0.000	19248.923
810	0.000	0.000	0.008	0.247	0.000	0.001	19257.279
811	0.000	0.000	0.000	0.113	0.000	0.000	19261.263
812	0.000	0.000	0.000	0.158	0.000	0.000	19269.842
813	0.000	0.000	0.001	0.231	0.000	0.000	19272.611
814	0.000	0.000	0.001	0.163	0.000	0.000	19274.805
815	0.000	0.000	0.003	0.124	0.000	0.000	19280.526
816	0.000	0.000	0.001	0.114	0.000	0.000	19284.114
817	0.000	0.000	0.003	0.046	0.000	0.000	19287.981
818	0.000	0.000	0.001	0.007	0.000	0.000	19289.965

Letting

$$\mathbf{C}_{n-1} = \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T \tag{A.3}$$

Hence Eq.A.3 becomes

$$\mathbf{B} = \mathbf{C}_{n-1} + \mathbf{x}_n \mathbf{x}_n^T = \mathbf{C}_n \tag{A.4}$$

The inverse of $\mathbf{B}$ is given by the Sherman-Morrison formula [32] as follows

$$\begin{aligned} \mathbf{B}^{-1} &= (\mathbf{C}_{n-1} + \mathbf{x}_n \mathbf{x}_n^T)^{-1} \\ &= \mathbf{C}_{n-1}^{-1} - \frac{\mathbf{C}_{n-1}^{-1} \mathbf{x}_n \mathbf{x}_n^T \mathbf{C}_{n-1}^{-1}}{1 + \mathbf{x}_n^T \mathbf{C}_{n-1}^{-1} \mathbf{x}_n} \end{aligned} \tag{A.5}$$

i.e.,

$$\mathbf{B}^{-1} = \mathbf{C}_n^{-1} = \mathbf{C}_{n-1}^{-1} - \frac{\mathbf{C}_{n-1}^{-1} \mathbf{x}_n \mathbf{x}_n^T \mathbf{C}_{n-1}^{-1}}{1 + \mathbf{x}_n^T \mathbf{C}_{n-1}^{-1} \mathbf{x}_n} \tag{A.6}$$

It is known that the Sherman-Morrison scheme may become numerically unstable owing to accumulative rounding errors resulting from the matrix-matrix subtractions [53, 54]. Mitigations for this problem are discussed in Section 4.

For reasons that will be later elobrated in Section 4 , let us slightly modify the above equation by multiplying ξ^2 on both sides of the equation, and then for notation clarity denote $\xi^2 \mathbf{C}^{-1}$ by $\mathbf{P}$.

$$\xi^2 \mathbf{B}^{-1} = \mathbf{P}_{n-1} - \frac{\mathbf{P}_{n-1} x_n x_n^T \mathbf{P}_{n-1}}{\xi^2 + \mathbf{x}_n^T \mathbf{P}_{n-1}^{-1} \mathbf{x}_n} \tag{A.7}$$

By induction, evidently, $\xi^2\mathbf{B}^{-1}$ can be given by a simple and initializable recursive algorithm, *i.e.*,

$$\xi^2\mathbf{B}^{-1} = \mathbf{P}_0 - \sum_{j=1}^{n} \frac{\mathbf{P}_{j-1}\mathbf{x}_j\mathbf{x}_j^T\mathbf{P}_{j-1}}{\varsigma_j} \tag{A.8}$$

where $\varsigma_j = \xi^2 + \mathbf{x}_j^T\mathbf{P}_{j-1}\mathbf{x}_j$; $\mathbf{P}_0$ encapsulates any *a priori* information we have about the physical system. We will come back to this, but for the time being (with no loss of generality) for clearity sake we set $\mathbf{P}_0$ to a unit matrix $\mathbf{I}_{n\times n}$.

Nevertheless, there is a problem with the algorithm: $\mathbf{P}$, a matrix of order n, is required in each iteration. In order to address this obvious core memory bottleneck, we express Eq. (A.8) in terms of vectors and scalars. This allows us to take advantage of the fact that in some judiously selected basis set [63, 64] the Hamiltonian matrix is very sparse — indeed the effectiveness of the scanning algorithms that were described in Section 2 rely on this very fact.

In the vector and scalar representation, the recursive algorithm is given as

$$\begin{aligned}
\mathbf{P}_1 &= \mathbf{I} - \frac{1}{\varsigma_1}\mathbf{x}_1^{2t} \\
\mathbf{P}_2 &= \mathbf{I} - \frac{1}{\varsigma_1}\mathbf{x}_1^{2t} - \frac{1}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}\mathbf{x}_1}{\varsigma_1})^{2t} \\
..... &= .... \\
\xi^2\mathbf{B}^{-1} &= \mathbf{I} - \frac{1}{\varsigma_1}\mathbf{x}_1^{2t} - \frac{1}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}\mathbf{x}_1}{\varsigma_1})^{2t} \\
&\quad - \frac{1}{\varsigma_3}\Big(\mathbf{x}_3 - \frac{\beta_{13}\mathbf{x}_1}{\varsigma_1} - \frac{\beta_{23}}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}\mathbf{x}_1}{\varsigma_1})\Big)^{2t} \\
&\quad \cdots \cdots \\
&\quad - \frac{1}{\varsigma_n}\Big\{\mathbf{x}_n - \frac{\beta_{1n}\mathbf{x}_1}{\varsigma_1} - \frac{\beta_{2n}}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}\mathbf{x}_1}{\varsigma_1}) \\
&\quad - \frac{\beta_{3n}}{\varsigma_3}\Big(\mathbf{x}_3 - \frac{\beta_{13}\mathbf{x}_1}{\varsigma_1} - \frac{\beta_{23}}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}\mathbf{x}_1}{\varsigma_1})\Big) - \dots \\
&\quad - \frac{\beta_{n-1n}}{\gamma_{n-1}}\Big[\mathbf{x}_{n-1} - \frac{\beta_{1n-1}\mathbf{x}_1}{\varsigma_1} - \frac{\beta_{2n-1}}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}\mathbf{x}_1}{\varsigma_1}) \\
&\quad - \frac{\beta_{3n-1}}{\varsigma_3}\Big(\mathbf{x}_3 - \frac{\beta_{13}\mathbf{x}_1}{\varsigma_1} - \frac{\beta_{23}}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}\mathbf{x}_1}{\varsigma_1})\Big) - \dots\Big]\Big\}^{2t}
\end{aligned} \tag{A.9}$$

where,
A^{2t} means AA^T;
$\beta_{1j} = \mathbf{x}_1^T\mathbf{x}_j$; $\beta_{2j} = (\mathbf{x}_2 - \frac{\beta_{12}}{\varsigma_1}\mathbf{x}_1)^T\mathbf{x}_j$; $\beta_{3j} = \Big(\mathbf{x}_3 - \frac{\beta_{13}}{\varsigma_1}\mathbf{x}_1 - \frac{\beta_{23}}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}}{\varsigma_1}\mathbf{x}_1)\Big)^T\mathbf{x}_j$; .....
$\varsigma_1 = \xi^2 + \mathbf{x}_1^T\mathbf{x}_1$; $\varsigma_2 = \xi^2 + \mathbf{x}_2^T(\mathbf{x_2} - \frac{\beta_{\mathbf{12}}}{\varsigma_{\mathbf{2}}}\mathbf{x_1})$; $\varsigma_3 = \xi^2 + \mathbf{x}_3^T\Big(\mathbf{x}_3 - \frac{\beta_{13}}{\varsigma_1}\mathbf{x}_1 - \frac{\beta_{23}}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}}{\varsigma_1}\mathbf{x}_1)\Big)$; .....

Let
$\mathbf{g}_1 = \mathbf{x}_1$
$\mathbf{g}_2 = \mathbf{x}_2 - \frac{\beta_{12}}{\varsigma_1}\mathbf{x}_1$
$\mathbf{g}_3 = \mathbf{x}_3 - \frac{\beta_{13}}{\varsigma_1}\mathbf{x}_1 - \frac{\beta_{23}}{\varsigma_2}(\mathbf{x}_2 - \frac{\beta_{12}}{\varsigma_1}\mathbf{x}_1)$;

....

$$\Rightarrow \mathbf{g}_j = \mathbf{x}_j - \sum_{i=1}^{j-1} \frac{\beta_{ij}}{\varsigma_i} \mathbf{g}_i \tag{A.10}$$

where $\varsigma_i = \xi^2 + \mathbf{x}_i^T \mathbf{g}_i$; $\beta_{ij} = \mathbf{g}_i^T \mathbf{x}_j$; j = 1,2, ...n. Note that $\mathbf{g}$ is a vector of order n, whereas β and ς are scalars.

Putting all this together, Eq. (A.10) becomes

$$\xi^2 \mathbf{B}^{-1} = \mathbf{I} - \sum_{j=1}^{n} \frac{\mathbf{g}_j \mathbf{g}_j^T}{\varsigma_j} \tag{A.11}$$

Obviously the number of terms on the RHS of Eq. (A.10) increase as j becomes bigger which creates CPU-time bottle-neck for calculating Eq. (A.11) when n is large.

$\mathbf{B}$ is symmetric positive definite, see Section 3, whose elements are given by

$$B_{ij} =< \mathbf{x}_i, \mathbf{x}_j >= \mathbf{x}_i^T \mathbf{x}_j \tag{A.12}$$

This means $\mathbf{B}$ can be associated with a Reproducing Kernel Hilbert Space (RKHS) where $\mathbf{x}_i s$ have representers that can span the RKHS [55, 56] in which $\mathbf{x}_i^T \mathbf{x}_j$ corresponds to a canonical dot product [57, 58], a linear Mercer kernel [59]. Moreover, according to Hilbert space theory [42], any pair of vectors in a Hilbert space satisfies the Cauchy-Schwartz inequality. In the RKHS case, $(\mathbf{x}_i^T \mathbf{x}_i)^{\frac{1}{2}} (\mathbf{x}_j^T \mathbf{x}_j)^{\frac{1}{2}} \geq \mid \mathbf{x}_i^T \mathbf{x}_j \mid$, *i.e.*,

$$\mid \mathbf{x}_i^T \mathbf{x}_i \mid > \frac{(\mathbf{x}_i^T \mathbf{x}_i)^{\frac{1}{2}}}{(\mathbf{x}_j^T \mathbf{x}_j)^{\frac{1}{2}}} \mid \mathbf{x}_i^T \mathbf{x}_j \mid \tag{A.13}$$

$$\Rightarrow \xi^2 + \mid \mathbf{x}_i^T \mathbf{x}_i \mid > \frac{(\mathbf{x}_i^T \mathbf{x}_i)^{\frac{1}{2}}}{(\mathbf{x}_j^T \mathbf{x}_j)^{\frac{1}{2}}} \mid \mathbf{x}_i^T \mathbf{x}_j \mid + \xi^2$$

$$\Rightarrow \frac{(\mathbf{x}_j^T \mathbf{x}_j)^{\frac{1}{2}}}{(\mathbf{x}_i^T \mathbf{x}_i)^{\frac{1}{2}}} \Big[1 - \frac{\xi_2}{\xi^2 + \mathbf{x}_i^T \mathbf{x}_i} \Big] > \mid \frac{\mathbf{x}_i^T \mathbf{x}_j}{\xi^2 + \mathbf{x}_i^T \mathbf{x}_i} \mid \tag{A.14}$$

The equality sign is dropped as $j > i$. With the appropriate value of ξ^2, $\mid \frac{\mathbf{x}_i^T \mathbf{x}_j}{\xi^2 + \mathbf{x}_i^T \mathbf{x}_i} \mid$<1. After some algebra, by the same token, it can also be deduced that $\mid \frac{\mathbf{g}_i^T \mathbf{x}_j}{\xi^2 + \mathbf{x}_i^T \mathbf{g}_i} \mid$<1.

Therefore only ($\lambda = j - i$) terms where i is close to j are significant for calculating $\mathbf{g}_j$ in Eq. (A.10): Since $\mid \frac{\mathbf{g}_i^T \mathbf{x}_j}{\xi^2 + \mathbf{x}_i^T \mathbf{g}_i} \mid = \mid \frac{\beta_{ij}}{\varsigma_i} \mid$, terms containing $\mid \frac{\beta_{ij}}{\varsigma_i} \mid$ or the products $\mid \frac{\beta_{kj} \beta_{mj} \beta_{\nu j} \cdots}{\varsigma_k \varsigma_m \varsigma_\nu \cdots} \mid \rightarrow 0$ quickly, where $\lambda < j - i$, can be dropped from the calculation rendering only few terms necessary to accurately calculate $\mathbf{g}_j$.

In principle, the value of λ can be estimated through Eq. (A.14), but in practice it can be chosen arbitrarily. In the latter case, one should bear in mind that λ is a lower bound for a given value of ξ^2. In principle, therefore, the pre-defined λ value must not be less than its lower limit for the given ξ^2, otherwise $\xi^2 \mathbf{B}^{-1}$ loses its positive definiteness and the Cauchy-Schwartz inequality does not hold. Nonetheless, since the Hamiltonian system is linear, (in practice) one can afford to ignore the positive definiteness lost as long as the eigenvalues of the few dominant eigenvalues of $\xi^2 \mathbf{B}^{-1}$ remain positive.

On this basis $\mathbf{g}_j = \mathbf{x}_j - \sum_{i=1}^{j-1} \frac{\beta_{ij}}{\varsigma_i}\mathbf{g}_i$ ($j = 1, 2, ..., n$ -1) can be divided into two parts

$$\mathbf{g}_j = \mathbf{x}_j - \begin{cases} \sum_{i=1}^{j-1} \frac{\beta_{ij}}{\varsigma_i}\mathbf{g}_i & \text{for } j-1 \leq \lambda \leq n-1 \\ \sum_{i=j-\lambda}^{j} \frac{\beta_{ij}}{\varsigma_i}\mathbf{g}_i & \text{for } \lambda < j \leq n \end{cases} \tag{A.15}$$

This alleviates the CPU-time bottle-neck problem.

In solving Eq. (12) constructing $\{\mathbf{g}_j\}_{j=1}^n$ once and storing the set would have been far more effecient computationally than generating the set for each matrix-vector multiplication step. However, storing a set of n dense vectors of size $n\times1$ is as difficult as keeping a $n\times n$ matrix in core-memory. Fortunately, as discussed before, for most physical systems the Hamiltonian matrix is sparse or diagonally dominant [63, 64]. This suggests that for a small λ, the set $\{\mathbf{g}_j\}_{j=1}^n$ is sparse as well. One can prove and quantify the level of $\{\mathbf{g}_j\}_{j=1}^n$ sparseness, but the mathematics is lengthy. Below we only motivate the proof.

In principle, $\mathbf{g}_j$ can be written as

$$\mathbf{g}_j = \begin{cases} \sum_{i=1}^{j-1} \gamma_i\mathbf{x}_i & \text{for } j-1 \leq \lambda \leq n-1 \\ \sum_{i=j-\lambda}^{j} \gamma_i\mathbf{x}_i & \text{for } \lambda < j \leq n \end{cases} \tag{A.16}$$

where $\gamma_i(\in \mathcal{R})$ is a collective index for all the $\frac{\beta_{ij}}{\varsigma_i}$'s coefficients of $\mathbf{x}_i$ in Eq. (A.15). For small λ, one easily sees that the sparsity of $\mathbf{g}_j$ is determined by the sparseness of $\mathbf{x}_i's$. Thus when the Hamiltonian matrix is sparse or dense (but diagonally dominant), $\mathbf{g}_j$ is sparse, as well, in absolute or relative terms: Absolute in the sense that some of the vector entries are already zeroes; relative in the sense that some values of the vector elements are negligible with respect to components with larger values. In the latter scenario, an arbitrary threshold ε can be pre-chosen, such that any value of $\mathbf{g}_j$ below ε is set to zero.

Appendix B. The rest of the proof in Section 2.1

B is (i) positive definite (ii) invertible (iii) diagonalizable.

First let us define what these terms mean:

- A square real symmetric matrix is positive definite if none of its eigenvalues $\leq$0.
- The matrix is invertible if it is nonsingular, *i.e.*, none of its eigenvalues is zero.
- It is diagonalizable if there is diagonal matrix similar to it.

(i) *Positive definiteness*: As $(\acute{\mathbf{H}} - \mathbf{I}\zeta)$ is nonsingular and real symmetric matrix, and $\mathbf{B}= (\acute{\mathbf{H}} - \mathbf{I}\zeta)^T(\acute{\mathbf{H}} - \mathbf{I}\zeta)$, then $\sigma_j(\mathbf{B}) > 0 \; \forall$ j

(ii) *invertibility*: $\mathbf{B}^{-1} = \left[(\acute{\mathbf{H}} - \mathbf{I}\zeta)(\acute{\mathbf{H}} - \mathbf{I}\zeta)\right]^{-1} \Rightarrow (\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1}(\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1}$. It has already been stated that $(\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-1}$ exists $\Rightarrow \mathbf{B}^{-1} = (\acute{\mathbf{H}} - \mathbf{I}\zeta)^{-2}$. Thus $\mathbf{B}$ is invertible.

(iii) *diagonalizability:* We know that $(\acute{\mathbf{H}} - \mathbf{I}\zeta)$ is diagonalizable — that is, $\mathbf{Z}^T(\acute{\mathbf{H}} - \mathbf{I}\zeta)\mathbf{Z} = \Lambda$ where Λ is a diagonal matrix $\Rightarrow \mathbf{B} = (\acute{\mathbf{H}} - \mathbf{I}\zeta)^2 = \mathbf{Z}\Lambda^2\mathbf{Z}^T \Rightarrow \mathbf{Z}^T\mathbf{B}\mathbf{Z} = \Lambda^2$. Hence $\mathbf{B}$ is diagonalizable.

[1] G. Hooft, 'In Search of the Ultimate Building Blocks',Cambridge University Press (1997).
[2] R. E. Wyatt, *Phys. Rev. E* **51**, 3643 (1995).
[3] H. G. Yu and G. Nyman, *J.Chem.Phys.* **110**, 11133 (1999).
[4] H. Y. Mussa and J. Tennyson, *Chem. Phys. Lett.* **366**, 449 (2002).
[5] U .V. Riss and H. D. Meyer,*J. Phys. B: At. Mol. Opt. Phys.* **26** 4503 (1993).
[6] D. Neuhauser,*J.Chem.Phys.* **93**, 2611 (1990).
[7] V. A. Mandelshtam and H. S. Taylor,**102**, 7390 (1995).
[8] R. Chen and H. Guo, *J. Chem. Phys.* **105**, 1311 (1996).
[9] H. G. Yu and S. C. Smith, *Chem. Phys. Lett.* **283**, 69 (1998).
[10] H. Zhang and S. C. Smith, *Phys.Chem.Chem.Phys.* **3**, 2282 (2001).
[11] R. Chen and H. Guo, *J. Comp. Phys.* **136**, 136 (1997).
[12] M. R. Wall and D. Neuhauser, *J.Chem.Phys.* **102**, 8011(1995).
[13] V. A. Mandelshtam and H. S. Taylor, *J. Chem. Phys.* **106**, 5085 (1997).
[14] V. A. Mandelshtam and H. S. Taylor, *J. Chem. Phys.* **107**, 6756 (1997).
[15] R. Chen and H. Guo, *J. Chem. Phys.* **111**, 464 (1999).
[16] H. Zhang and S. C. Smith, *Chem. Phys. Lett.* **347**, 211 (2001).
[17] M. Alacid, C. Leforestiera and N. Moiseyev, *Chem. Phys. Lett.* **305**, 258(1999).
[18] W. Zhu, Y. Huang, D. J. Kouri, C. Chandler and D. K. Hoffman, *Chem. Phys. Lett.* **217**, 73 (1994).
[19] D. J. Kouri, W. Zhu, G. A. Parker and D. K. Hoffman, *Chem. Phys. Lett.* **238**, 395 (1995).
[20] T. Ericsson and A. Ruhe, *Math. Comput.*, **35**, 1251 (1980).
[21] S. Huang and T. Carrington Jr., *J. Chem. Phys.* **112**, 8765 (2000).
[22] C. Leforestier, K. Yamashita and N. Moiseyev, *J. Chem. Phys.* **103**, 8468 (1995).
[23] B. Poirier and T. Carrington Jr., *J. Chem. Phys.* **116**, 1215 (2001).
[24] S. Dallwig, N. Fahrer, and C. Schlier, *Chem. Phys. Lett.* **191**, 69 (1992).
[25] L. W. Wang and A. Zunger, *J. Chem. Phys.* **100**, 2394 (1994).
[26] D. A. Drabold, *J.Non. Cryst. Sol.* **266**, 211 (2000).
[27] H. De Raedt and W. von der Linden, *Phys. Rev.* **B45**, 8787 (1992)
[28] H. De Raedt and M. Frick, *Phys. Rep.* **231**, 107 (1993).
[29] J. W. Clark, D. G. Lucarelli and T. Tarn,*Int. J. Mod. Phys.* **B17**, 5397 (2003).
[30] D. S. Watkins, 'Fundamentals of Matrix Computation', Wiley (1991)
[31] S. Pissanetzky, 'Sparse Matrix Technology', MIR, Moscow (1988).
[32] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, 'Numerical Recipes in Fortran: The Art of Scientific Computing', Cambridge University Press (1992)
[33] Y. Saad,'Numerical Methods For Large Eigenvalue Problems((Algorithms and Architecture for Advanced Scientific Computing)', Halsted Press(1992).
[34] L .I. Shiff, 'Quantum Mechanics', McGraw-Hill (1968).
[35] M. Scheter, 'Opertor Methods in Quantum Mechanics', North-Holland (1981).
[36] C. Lanczos, *J. Res. Natl. Bur. Stand.* **45**, 255 (1950).
[37] J. Cullum and R. A.Willoughby,*J.comp.Phys.* **44**, 329 (1981).
[38] H. A. van der Vorest, 'Iterative Krylov Methods for Large Linear Systems', Cambridge University Press (2003)
[39] T. C. Oppe, W. D. Joubert, and D. C. Kincaid, *Comput.Phys.Commun.* **53**,283 (1989).
[40] C. Iung and C. Leforestier, *J. Chem. Phys.* **102**, 8453 (1995).

[41] H. Kono, *Chem. Phys. Lett.* **214**, 137 (1993).
[42] C. J. Isham, 'Lectures on Quantum Theory: mathematical and structural foundations', Imperial College Press (1995).
[43] R. Clifton, 'Introductory Notes on the Mathematics Needed for Quantum Theory', 'http://philsci-archive.pitt.edu/archive/00000390/'
[44] S. Skokov, J. Qi, J. M. Bowman, C. Yang, S. K. Gray, K. A. Peterson and V. A. Mandelshtam,*J. Chem. Phys.* **109**, 10273(1998).
[45] S. Huang and T. Carrington Jr.,*Chem.Phys.Lett.* **312**,311 (1999).
[46] H. Zhang and S. C. Smith,*Chem.Phys.Lett.* **347**,211 (2001).
[47] R. Chen and H. Guo, *J. Chem. Phys.* **108**, 6068 (1998).
[48] G. Ma and H. Guo, *J. Chem. Phys.* **111**, 4032 (1999).
[49] Y. L. Cun, I. Kanter, and S. A. Solla, *Phys. Rev. Lett.* **66**, 2396 (1991).
[50] C. Wunsch, 'The Ocean Circulation Inverse Problem', Chapter3, Cambridge University Press (1991)
[51] Available at *http://www.caam.rice.edu/software/ARPACK/*
[52] 'EA22: Sparse symmetric: simultaneous iteration' at *http://hsl.rl.ac.uk/contentshsl2004.html*
[53] S. Haykin, *Neural Networks: A Comprehensive Foundation (2nd edition)*, Prentice Hall (1998).
[54] G. Bierman, *Factorization Methods for Discrete Sequential Estimation*, Academic Press (**1977**).
[55] N. Aronszajn, *Trans. Amer. Math. Soc.* **68**, 337 (1950).
[56] J. M. Moguerza and A. Muoz,*Statistical Science* **21**, 322(2006).
[57] R. Rosipal and L. J. Trejo, *J. Machine Learning Res.* **2**, 97 (2001).
[58] B. Scholkopf, S. Mika, C. J. Burges, P. Knirsch, K. Muller, G. Ratsch and A. J. Smola, *IEEE Trans. on Neural Networks* **10**, 1000 (1999).
[59] J. Mercer, *Pholos. Trans.Roy. Soc. London* **209**, 415 (1909).
[60] H. van Storch and F. W. Zwiewg, 'Statistical Analysis in Climate Research', Cambridge University Press (2001).
[61] C. d. Rodgers,'Inverse Methods For Atmospheric Sounding: theory and practice ', World Scientific (2000).
[62] F. M. Ham and I. Kostanic, 'Principles of Neurocomputing for Sciences and Engineering', McGraw-Hill (2000).
[63] J. C. Light, I. P. Hamilton, and J. V. Lill, *J. Chem. Phys.* **82**, 1400 (1985).
[64] J. Tennyson, M. A. Kostin, P. Barletta, G. J. Harris, J. Ramanlal, O. L. Polyansky and N. F. Zobov, *Computer Phys. Comm.*, **163**, 85 (2004).
[65] H. Y. Mussa, J. Tennyson, and R. C. Glen, *J. Chem. Phys.* (to be submitted).
[66] J. Tennyson and B. T. Sutcliffe, *J. Chem. Phys.*, **77**, 4061 (1982).
[67] S. Skokov, J. Qi, J. M. Bowman, and K. A. Peterson, *J. Chem. Phys.*, **109**, 2662 (1998).
[68] Subroutine F02ABF, National Algorithms Group (NAG) (2002), see *www.nag.co.uk/nagware/mt/doc/f02abf.html*.