Effective interactions and convergence criteria for configuration interaction methods

Morten Hjorth-Jensen[1,2] and Simen Kvaal[2]

[1]Department of Physics, University of Oslo, N-0316 Oslo, Norway

[2]Center of Mathematics for Applications, University of Oslo, N-0316 Oslo, Norway

April 30, 2009

Outline

Motivation

1. Can we find a similarity transformation which reduces the role of many-body forces?
2. Which conditions give such an optimal reduction?
3. How do we define good model spaces? And which consequences does a good model space have for many-body interactions?
4. How do we link standard many-body methods and EFT?

In this talk I will first discuss some mathematical properties of effective interactions and CI methods applied to quantum dots. Then I will show a model many-body problem where many-body forces and the definition of the model space play significant roles.

Definitions

We consider a general A-body situation where a Hilbert space of finite dimension n is given along with an A-body Hamiltonian $\hat{H}$ with spectral decomposition given by

$$\hat{H} = \sum_{k=1}^{n} E_k \left|\psi_k\right\rangle \left\langle\psi_k\right|,$$

where $\{|\psi_k\rangle\}_{k=1}^{n}$ is an orthonormal basis of eigenvectors and where E_k are the corresponding eigenvalues. We choose the dimension n to be finite for simplicity, but the theory may be generalized to infinite dimensional settings where $\hat{H}$ has a purely discrete spectrum.

Definitions

The Hilbert space is divided into the model space $\mathcal{P}$, and its complement, denoted by $\mathcal{Q}$. We assume that $m = \dim(\mathcal{P})$.
The effective Hamiltonian $\hat{H}_{\text{eff}}$ is defined in the $\mathcal{P}$-space *only*, and by definition its eigenvalues are identical to the m eigenvalues of $\hat{H}$. This is equivalent to $\hat{H}_{\text{eff}}$ being given by

$$\begin{aligned} \hat{H}_{\text{eff}} &:= \hat{P}\mathcal{H}\hat{P} \\ &= \hat{P}e^{-\hat{G}}\hat{H}e^{\hat{G}}\hat{P}, \end{aligned}$$

where $\mathcal{H}$ is assumed to obey the de-coupling equation

$$\hat{Q}\mathcal{H}\hat{P} = 0.$$

If the latter is satisfied, the $\mathcal{P}$-space is easily seen to be invariant under $\mathcal{H}$, and since similarity transformations preserve eigenvalues, $\hat{H}_{\text{eff}}$ is seen to have m eigenvalues of $\hat{H}$.

Definitions

Without loss of generality, we assume that the eigenvalues E_k of $\hat{H}$ are arranged so that $\hat{H}_{\text{eff}}$, which is non-Hermitian in general, has the spectral decomposition

$$\hat{H}_{\text{eff}} = \sum_{k=1}^{m} E_k \, |\phi_k\rangle \, \langle\tilde{\phi}_k|,$$

where $\{|\phi_k\rangle\}_{k=1}^{m}$ is a basis for the $\mathcal{P}$-space, and where $\langle\phi_k|\tilde{\phi}_\ell\rangle = \delta_{k,\ell}$ defines the bi-orthogonal basis $\{|\tilde{\phi}_k\rangle\}_{k=1}^{m}$.

Definitions

The similarity transform operator $\exp(\hat{G})$ is of course not unique; E_k, $k = 1, \cdots, m$ can be chosen in many ways, and even if the effective eigenvector $|\phi_k\rangle$ is chosen to be related to $|\psi_k\rangle$, there is still great freedom of choice left.

Assume that we have determined the eigenvalues E_k, $k = 1, \ldots, m$ that $\hat{H}_{\text{eff}}$ should have. Two choices of the corresponding $|\phi_k\rangle$ are common: The Bloch-Brandow choice, and the canonical Van Vleck choice, resulting in "the non-Hermitian" and "the Hermitian" effective Hamiltonians, respectively. For a discussion of these approaches see the recent work of Simen Kvaal, Phys. Rev. C **78**, 044330 (2008).

Definitions

In the Bloch-Brandow scheme, the effective eigenvectors are simply chosen as

$$|\phi_k\rangle := \hat{P}\,|\psi_k\rangle\,,$$

which gives meaning whenever $\hat{P}\,|\psi_k\rangle$ defines a basis for $\mathcal{P}$-space. In this case, $\hat{G} = \hat{\omega}$, where $\hat{\omega} = \hat{Q}\hat{\omega}\hat{P}$, defined by

$$\hat{\omega}\hat{P}\,|\psi_k\rangle := \hat{Q}\,|\psi_k\rangle\,, \quad k = 1, \cdots, m.$$

Definitions

In contrast, the canonical Van Vleck effective Hamiltonian chooses a certain orthogonalization of $\{\hat{P}\,|\psi_k\rangle\}_{k=1}^{m}$ as effective eigenvectors. In this case, $\hat{G} = \text{artanh}(\hat{\omega} - \hat{\omega}^{\dagger})$, which relates the two effective Hamiltonians to each other. The canonical effective interaction $\hat{H}_{\text{eff}}$ minimizes the quantity Δ defined by

$$\Delta(|\chi_1\rangle, \cdots, |\chi_m\rangle) := \sum_{k=1}^{m} \| \, |\chi_k\rangle - |\psi_k\rangle \, \|^2, \tag{1}$$

where the minimum is taken with respect to all orthonormal sets of $\mathcal{P}$-space vectors χ. The Bloch-Brandow effective eigenvectors, on the other hand, yield the *global* minimum of Δ.

Definitions

We have not specified *which* of the eigenvalues of H is to be reproduced by $\hat{H}_{\text{eff}}$. In general, we would like it to reproduce the ground state and the other lowest eigenstates of H if $m > 1$.

We define the effective *interaction* $\hat{V}_{\text{eff}}$ as

$$\hat{V}_{\text{eff}} := \hat{H}_{\text{eff}} - \hat{P}\hat{H}_0\hat{P},$$

where $[\hat{H}_0, \hat{P}] = 0$ is assumed. This is satisfied whenever the model space is spanned by Slater determinants being eigenvectors of $\hat{H}_0$. Common choice in many-body physics.

Algorithm

Let us briefly summarize the numerical algorithm for computing the (sub-cluster) effective a-body interaction $\hat{V}_{\text{eff}}^{(a)}$. (see Phys. Rev. C **78**, 044330 (2008)).
We assume here, that the computational basis is given by $\{|k\rangle\}_{k=1}^{n}$, and that $\hat{P}$ and $\hat{Q}$ are given by

$$\hat{P} \equiv \sum_{k=1}^{m} |k\rangle \langle k|$$

and

$$\hat{Q} \equiv \hat{\mathbf{1}} - \hat{P} = \sum_{k=m+1}^{n} |k\rangle \langle k| ,$$

respectively. We will let all operators be represented by their matrices in this basis.

Algorithm

The numerical algorithm can then be summarized as follows:

- Completely diagonalize the a-body Hamiltonian matrix H, viz,

$$H = UEU^{\dagger},$$

  where U is an $n \times n$ unitary matrix, where $U_{j,k} = \langle j|\psi_k\rangle$, and where E is diagonal, with $E_{k,k} = E_k$.

- Sort the eigenvalues and eigenvectors, i.e., permute the eigenvalues and columns of U, according to increasing energy or model space overlap, whatever is desired.
- Let $\tilde{U} = PUP = U_{1\ldots m,1\ldots m}$, i.e., the upper $m \times m$ block.

Algorithm

- Compute the singular value decomposition (SVD) of $\tilde{U}$, viz,

$$\tilde{U} = X\Sigma Y^{\dagger},$$

where X and Y are unitary matrices, and Σ is diagonal with diagonal elements $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_m$.

- Compute the $m \times m$ matrix V given by

$$V = XY^{\dagger}, \tag{2}$$

with $V_{j,k} = \langle j|\phi_k\rangle$. Eq. (2) solves the minimization of Δ in Eq. (1).

- Compute the effective interaction matrix, viz,

$$V_{\text{eff}} = VE_{1\ldots m,1\ldots m}V^{\dagger} - PH_0P.$$

- Extract occupation number matrix elements $u^{\alpha_1,\cdots,\alpha_a}_{\beta_1,\cdots,\beta_a}$ from V_{eff}.

Definitions and convergence criteria

The Hamiltonian of the quantum dot is given by

$$H := T + U, \tag{3}$$

where T is the many-body HO Hamiltonian, and U is the inter-electron Coulomb interactions. Thus, in dimensionless units,

$$U := \sum_{i<j}^{N} C(i,j) = \sum_{i<j}^{N} \frac{\lambda}{\|\vec{r}_i - \vec{r}_j\|}.$$

The parameter λ measures the strength of the interaction over the confinement of the HO, viz,

$$\lambda := \frac{1}{\hbar\omega}\left(\frac{e^2}{4\pi\epsilon_0\epsilon}\right),$$

where we recall that $\sqrt{\hbar/m\omega}$ is the length unit. Typical values for GaAs semiconductors are close to $\lambda = 2$. Increasing the trap size leads to a larger λ, and the quantum dot then approaches the classical regime.

Definitions and convergence criteria

For two particles in a parabolic trap, for certain values of the trap, one has closed-form solutions (Taut 1993 and 1994).
The Hamiltonian (3) becomes

$$H = -\frac{1}{2}(\nabla_1^2 + \nabla_2^2) + \frac{1}{2}(r_1^2 + r_2^2) + \frac{\lambda}{r_{12}}, \tag{4}$$

where $r_{12} = \|\vec{r}_1 - \vec{r}_2\|$ and $r_j = \|\vec{r}_j\|$. Introduce a set of scaled centre of mass coordinates given by $\vec{R} = (\vec{r}_1 + \vec{r}_2)/\sqrt{2}$ and $\vec{r} = (\vec{r}_1 - \vec{r}_2)/\sqrt{2}$.
This leads to the separable Hamiltonian

$$\begin{aligned} H &= -\frac{1}{2}(\nabla_r^2 + \nabla_R^2) + \frac{1}{2}(\|\vec{r}\|^2 + \|\vec{R}\|^2) + \frac{\lambda}{\sqrt{2}\|\vec{r}\|} \\ &= H_{\mathrm{HO}}(\vec{R}) + H_{\mathrm{rel}}(\vec{r}). \end{aligned}$$

Definitions and convergence criteria

The ground state (having $m = m' = 0$, $n = n' = 0$) for $\lambda = 1$ is given by

$$\Psi_0(\vec{R}, \vec{r}) \quad = \quad \frac{D}{\sqrt{2}}(r_{12} + \sqrt{2}a)e^{-(r_1^2+r_2^2)/2},$$

with D and a being constants.

Observe that this function has a cusp at $r = 0$, i.e., at the origin $x = y = 0$ (where we have introduced Cartesian coordinates $\vec{r} = (x, y)$ for the relative coordinate). Indeed, the partial derivatives $\partial_x\psi_0$ and $\partial_y\psi_0$ are not continuous there. The cusp stems from the famous "cusp condition" which simply states that, for a non-vanishing wave function at $r_{12} = 0$, the Coulomb divergence must be compensated by a similar divergence in the Laplacian.

Definitions and convergence criteria

On the other hand, the non-smooth function $\Psi_{0,0}(\vec{R},\vec{r})$ is to be expanded in the HO eigenfunctions, e.g., Fock-Darwin orbitals. For $m=0$, we have

$$\Phi^{\rm FD}_{n,0}(r) = \sqrt{\frac{2}{\pi}} L_n(r^2) e^{-r^2/2},$$

using the fact that these are independent of θ. Thus,

$$\Psi_0(\vec{r}) = \Phi^{\rm FD}_{0,0}(R) u_{0,0}(r) = \Phi^{\rm FD}_{0,0}(R) \sum_{n=0}^{\infty} c_n \Phi^{\rm FD}_{n,0}(r), \tag{5}$$

The functions $\Phi^{\rm FD}_{n,0}(r)$ are very smooth, as is seen by noting that $L_n(r^2) = L_n(x^2+y^2)$ is a polynomial in x and y, while $u_{0,0}(r) = u_{0,0}(\sqrt{x^2+y^2})$, so Eqn. (5) is basically approximating a square root with a polynomial. Therefore, the cusp at $r=0$ cannot be well approximated.

Definitions and convergence criteria

One can show that the smoothness properties of the wave function Ψ is equivalent to a certain decay rate of the coefficients c_n in Eqn. (5) as $n \longrightarrow \infty$. In this case, we will show that

$$\sum_{n=0}^{\infty} n^k |c_n|^2 < +\infty,$$

so that

$$|c_n| = o(n^{-(k+1+\epsilon)/2}). \tag{6}$$

Here, k is the number of times Ψ may be differentiated weakly, i.e., $\Psi \in H^k(\mathbb{R}^2)$, and $\epsilon \in [0, 1)$ is a constant. For the function Ψ_0 we have $k = 1$. This kind of estimate directly tells us that an approximation using only a few HO eigenfunctions necessarily will give an error depending directly on the smoothness k.

Definitions and convergence criteria

In a many-body context, define all possible Slater determinants for A particles in d dimensions, constrained by the energy

$$E \leq R_{ho} + \frac{Ad}{2}$$

where R_{ho} is the last oscillator shell, one can show (see Simen Kvaal, arxiv:0808.2145, and Phys. Rev. B, in press) that the error in the many-body energy by omitting shells above R_{ho} is given by

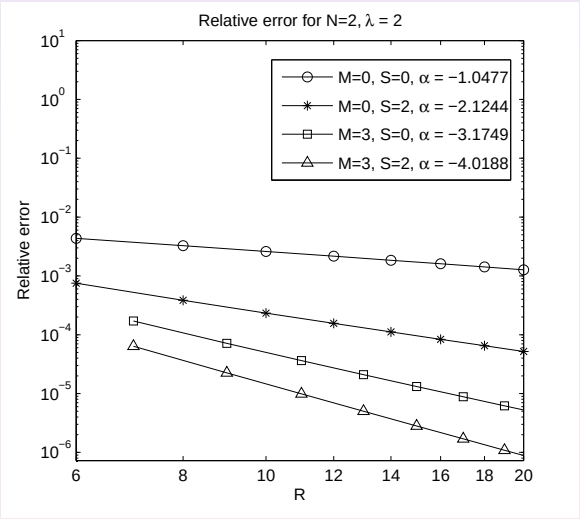
$$\Delta E \leq C \sum_{n=R_{ho}+1}^{\infty} (n + \frac{1}{2})|c_n|^2,$$

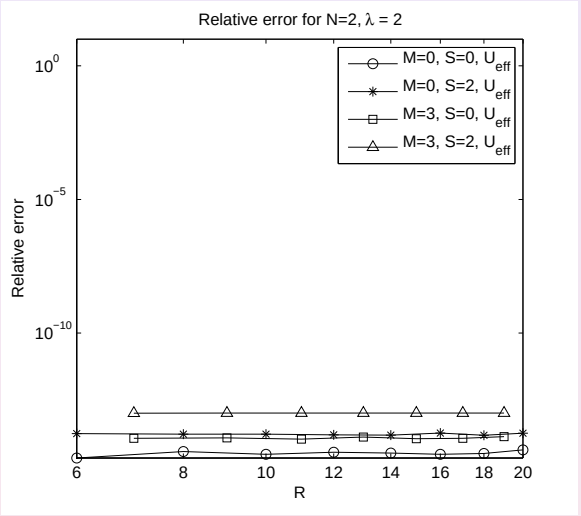
with C a constant and replacing the sum by an intergral it can be approximated as

$$\Delta E \sim O(R_{ho}^{-k+\epsilon-1}).$$

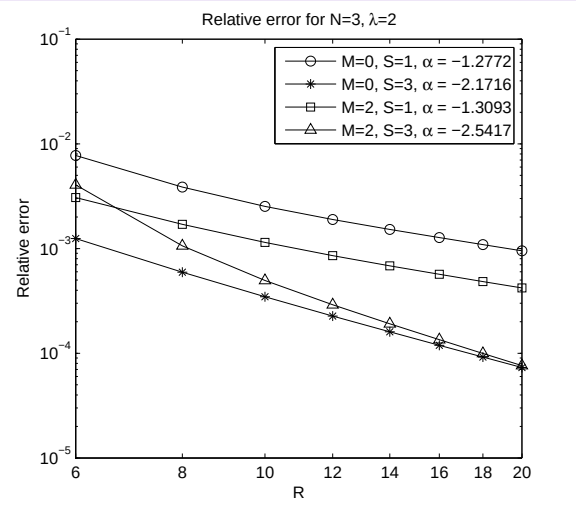
This gives a precise error estimate when truncating the many-particle Hilbert space. It depends on the properties (cusp conditions) of the wave function at $r = 0$.

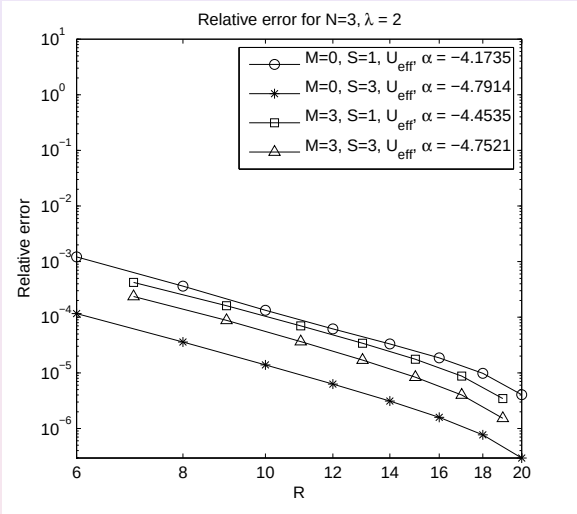
Numerical experiments, bare interaction vs effective



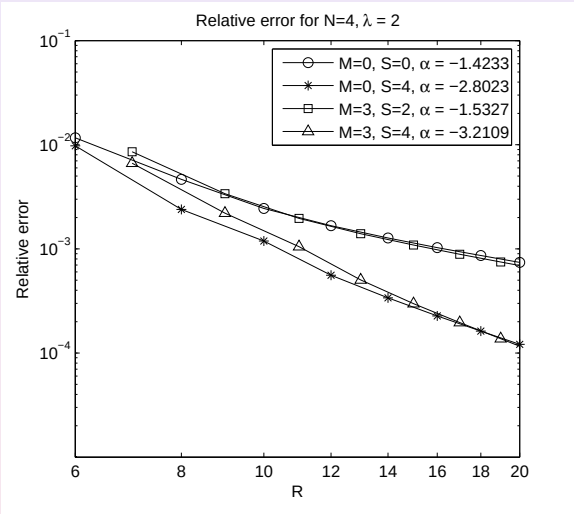


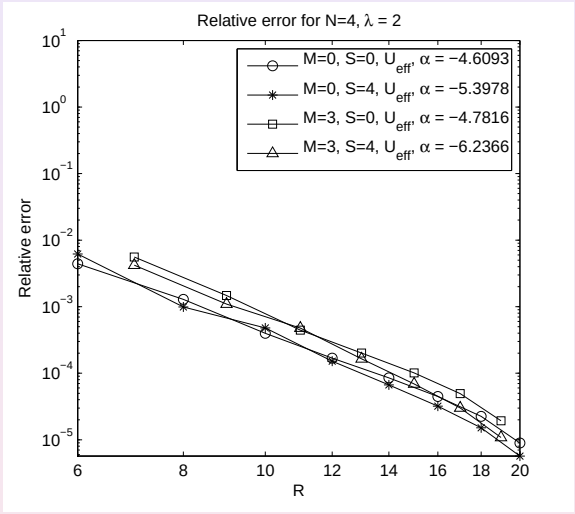
Numerical experiments, bare interaction vs effective



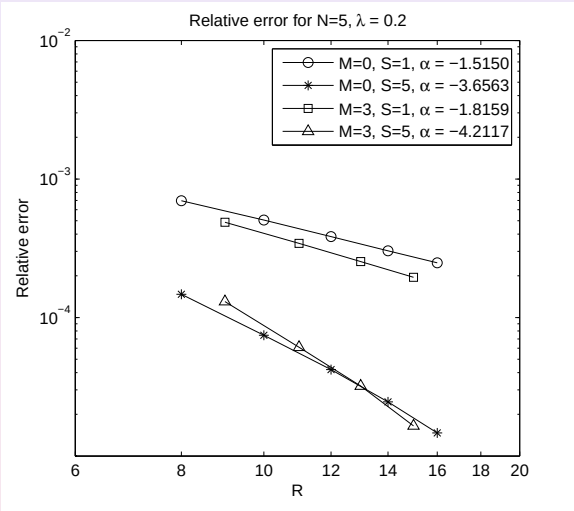


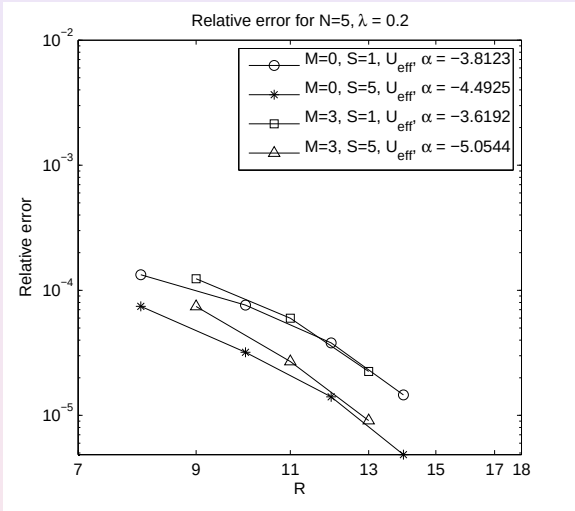
Numerical experiments, bare vs effective interaction





Numerical experiments, bare vs effective interaction





Petit summary

- Quantum dots are in general kind problems. The two-particle problem has a closed-form solution.
- Precise convergence criterion for bare interaction
- An effective interaction improves considerably the convergence, and a much small set of single-particle states is necessary. However, not easy to find precise convergence criterion
- Coupled-cluster calculations should exhibit similar convergence patterns
- Can extend the convergence criterion to other basis functions
- Can we define a priori what is a good model space?

A nastier case

Our specific model consists of N doubly-degenerate and equally spaced single-particle levels labelled by $p = 1, 2, \ldots$ and spin $\sigma = \pm 1$.
We write the Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{V},$$

where

$$\hat{H}_0 = \xi \sum_{p\sigma} (p-1) a^\dagger_{p\sigma} a_{p\sigma}$$

and

$$\hat{V} = -\frac{1}{2} g \sum_{pq} a^\dagger_{p+} a^\dagger_{p-} a_{q-} a_{p+} - \frac{1}{2} f \sum_{pqr} \left(a^\dagger_{p+} a^\dagger_{p-} a_{q-} a_{r+} + \text{h.c.} \right)$$

A nastier case

Here, H_0 is the unperturbed Hamiltonian with a spacing between successive single-particle states given by ξ, which we may set to a constant value $\xi = 1$ without loss of generality. The two-body operator $\hat{V}$ has two terms. The first term represents the pairing contribution and carries a constant strength g (it easy to extend our model to include a state dependent interaction). The indices $\sigma = \pm$ represent the two possible spin values.

A nastier case

The first term of the interaction can only couple pairs and excites therefore only two particles at the time.
The second interaction term, carrying a constant strength f, acts between a set of particles with opposite spins and allows for the breaking of a pair or just to excite a single-particle state. The spin of a given single-particle state is not changed. This interaction can be interpreted as a particle-hole interaction if we label single-particle states within the model space as hole-states. The single-particle states outside the model space are then particle states.

In our model we have kept both the interaction strength and the single-particle level as constants. In a realistic system like a nucleus this is not the case, however if a harmonic oscillator basis is used, as done in the no-core shell-model calculations at least the single-particle basis mimicks the input to realistic calculations.

A nastier case

Finding $\hat{H}_{\text{eff}}$ is equivalent to solving the original problem. In order to be useful, we need some sort of approximation scheme to find $\hat{H}_{\text{eff}}$.
In our model, $\mathcal{P}$ is defined by restricting the allowed leves accessible for the A particles under study. Thus,

$$\mathcal{P}^{(A)} := \text{span}\left\{|(p_1,\sigma_1)\cdots(p_A,\sigma_A)\rangle \;:\; p_k \leq N_P\right\},$$

where $N_P \leq N$ is the number of levels accessible in the model space. This is the way model spaces in general are built up, simply restricting the single-particle orbitals accessible. It is by no means the only possible choice. On the other hand, this way of defining the model space has a very intuitive appeal, as it naturally leads to a view of $\hat{H}_{\text{eff}}$ as a renormalization of $\hat{H}$. It also gives a natural relation between model spaces for different A, which is absolutely necessary for the sub-cluster effective Hamiltonian to be meaningful.

A nastier case

The effective Hamiltonian is seen to be an A-body operator in general, even though H itself may contain only two-body operators. Thus, $\hat{V}_{\text{eff}}$ can be written in its most general form as

$$\hat{V}_{\text{eff}}^{(A)} = \sum_{\alpha_1,\cdots,\alpha_A} \sum_{\beta_1,\cdots,\beta_A} u_{\beta_1,\cdots}^{\alpha_1,\cdots} a_{\alpha_1}^\dagger \cdots a_{\alpha_A}^\dagger a_{\beta_A} \cdots a_{\beta_1},$$

where $\alpha_k = (p_k, \sigma_k)$ and $u_{\beta_1,\cdots}^{\alpha_1,\cdots}$ represents a specific matrix element. The approximation idea is then to obtain instead an a-body effective interaction $\hat{V}_{\text{eff}}^{(a)}$, where $a < A$, and view this as an approximation to $\hat{V}_{\text{eff}}^{(A)}$. This leads to

$$\hat{V}_{\text{eff}}^{(A)} \approx \frac{\binom{A}{2}}{\binom{A}{a}\binom{a}{2}} \hat{V}_{\text{eff}}^{(a)},$$

which is a much simpler operator, usually obtainable exactly by large-scale diagonalization of the a-body Hamiltonian.

A nastier case

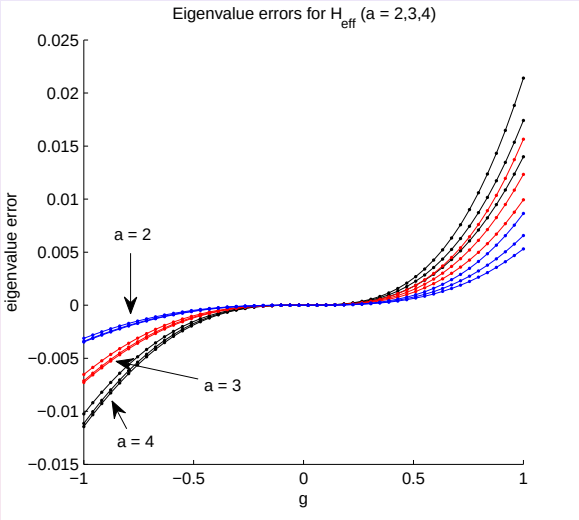
The remaining question is which eigenpairs of $H^{(a)}$ should be reproduced by $\hat{H}^{(a)}_{\text{eff}}$, and which approximate eigenvectors should be used. There is no unique answer to this. The "best" answer would for each problem require a complete knowledge of the conserved observables of the many-body Hamiltonian.

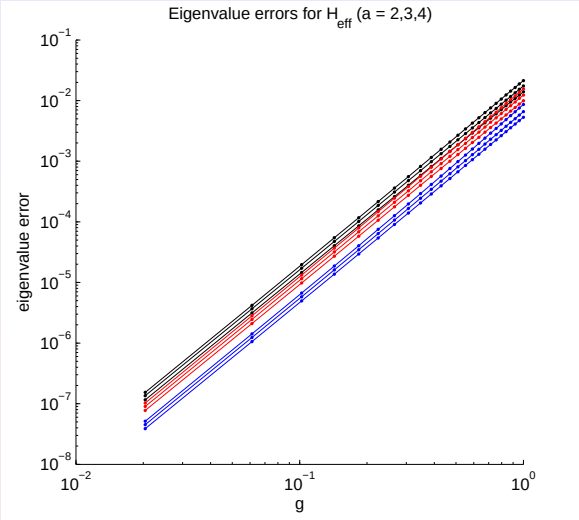
On the other hand, if $\hat{V}$ is a small perturbation, that is, we let $\hat{V} \mapsto \lambda \hat{V}$ and consider an adiabatic turning on by slowly increasing λ, then it is natural to choose the eigenvalues developing adiabatically from $\lambda = 0$. Indeed, $\hat{V}^{(a)}_{\text{eff}}$ is then seen to be identical to a class of a-body terms in the perturbation series for the full $\hat{V}^{(A)}_{\text{eff}}$ to infinite order. The problem is, that there is no way in general to decide *which* eigenvalues have developed adiabatically from $\lambda = 0$, and we must resort to a heuristic procedure.

A nastier case

Two alternatives present themselves as obvious candidates: Selecting the smallest eigenvalues, and selecting the eigenvalues whose eigenvectors have the largest overlap $\langle\psi_k|P|\psi_k\rangle$ with $\mathcal{P}$. Both are equivalent for sufficiently small λ, but the eigenvalues will cross in the presence of so-called intruder states for larger λ. Moreover, the presence of perhaps unknown constants of motion will make the selection by eigenvalue problematic, as exact crossings may lead us to select eigenpairs with $\hat{P}\,|\psi_k\rangle = 0$, which makes $\hat{H}^{(a)}_{\text{eff}}$ ill-defined. We therefore consider selection by model space overlap to be more robust in general.

A nastier case, numerical experiments, only pairing

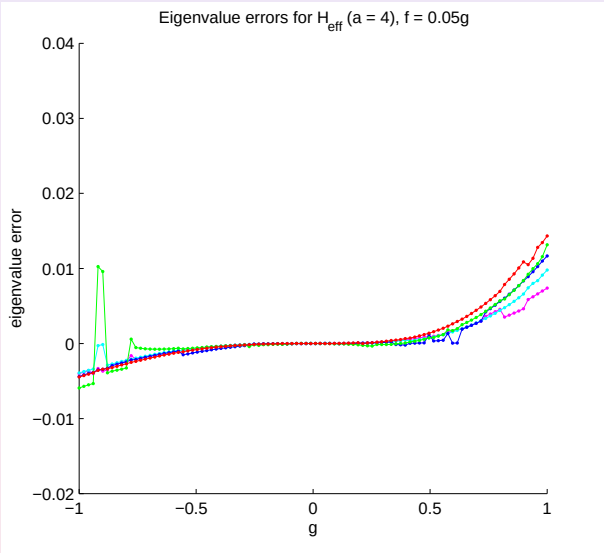


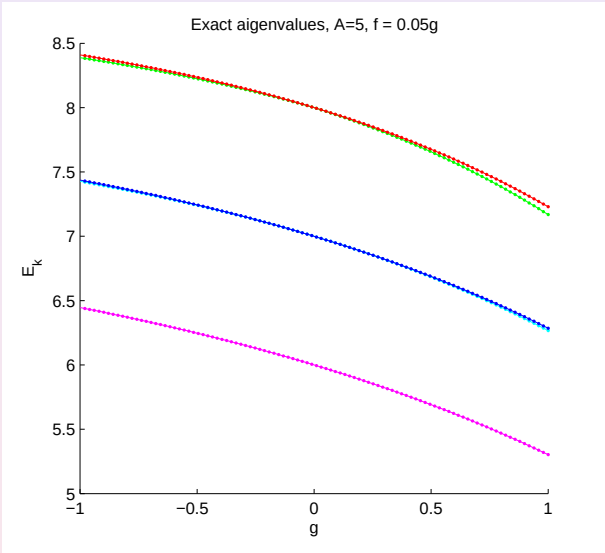


A nastier case, numerical experiments, only pairing

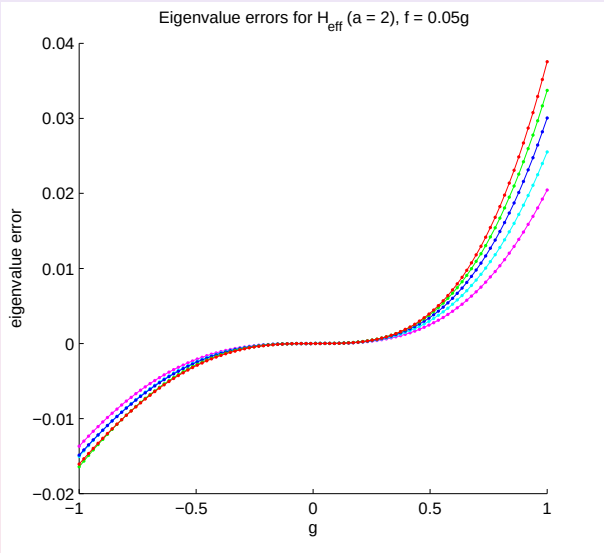
For $g > 0$, a double-logarithmic plot reveals an almost perfect g^3-behaviour of all errors. Larger values of a give smaller errors, as one would expect, but only by a constant factor. Thus, all the $\hat{V}_{\text{eff}}^{(a)}$ seem to be equivalent to perturbation theory to second order in the strength g with respect to accuracy. This order is constant, even though the complexity of calculating $\hat{V}_{\text{eff}}^{(a)}$ increases by orders of magnitude. Most of the physical correlations are thus well-represented by a two-body effective interaction. This is expected since a pairing-type interaction favors strong two-particle clusters. The choice of a constant pairing strength enhances also this type of correlations. Three-body and four-body clusters tend to be small.

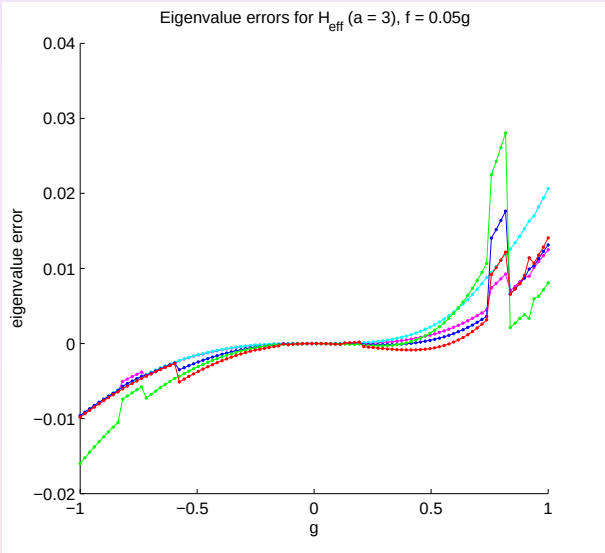
A nastier case, numerical experiments, particle-hole as well, four and five particles



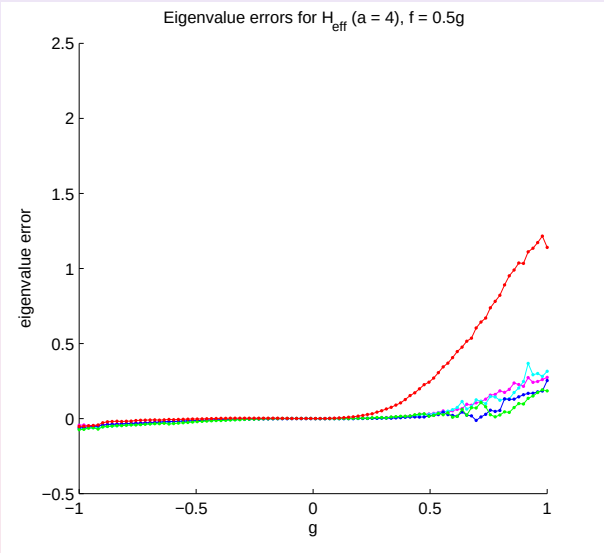


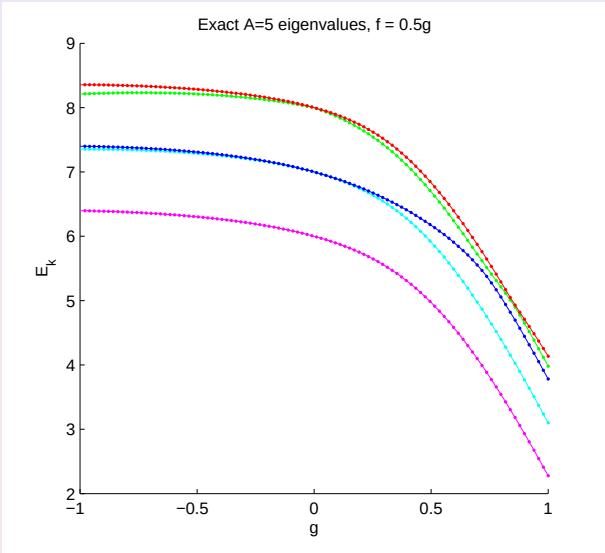
A nastier case, numerical experiments, particle-hole as well, two and three particles



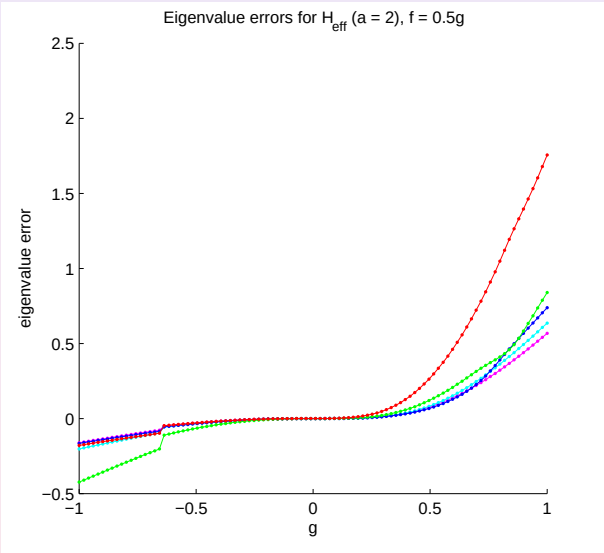


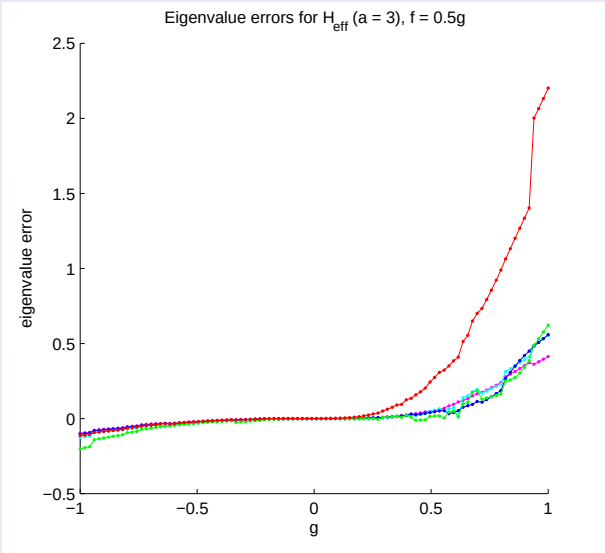
A nastier case, numerical experiments, particle-hole as well, four and five particles





A nastier case, numerical experiments, particle-hole as well, two and three particles





Conclusions

- Precise convergence criterion for bare Coulomb interactions
- An effective interaction improves considerably the convergence, and a much small set of single-particle states is necessary. However, not easy to find precise convergence criterion.
- Coupled-cluster calculations for Coulomb problems should exhibit the same pattern. What about nuclear forces? EFT-many-body strategy?
- When strong particle-hole correlations are present one may have problems in defining a proper model space. Level crossings and intruder states (Island of inversion problems in nuclear physics)
- We cannot define a priori what is a good model space and many-body interactions beyond two-body may cause problems.
- How can we link such problems with EFT?