Variations on the 'exact factorization' theme

Xavier Gonze[1,4], Jianqiang Sky Zhou[2,3,4], and Lucia Reining[2,4]

[1] Institute of Condensed Matter and Nanosciences, Université catholique de Louvain, Chemin des étoiles 8, bte L07.03.01, B-1348 Louvain-la-Neuve, Belgium
[2] Laboratoire des Solides Irradiés, École Polytechnique, CNRS, CEA, Université Paris-Saclay, F-91128 Palaiseau, France
[3] Sorbonne Universitè, CNRS, Institut des Nanosciences de Paris, UMR7588, F-75252, Paris, France
[4] European Theoretical Spectroscopy Facility (ETSF)

Abstract In a series of publications, Hardy Gross and co-workers have highlighted the interest of an "exact factorization" approach to the interacting electron-nuclei problem, be it time-independent or time-dependent. In this approach an effective potential governs the dynamics of the nuclei such that the resulting N-body nuclear density is in principle exact. This contrasts with the more usual adiabatic approach, where the effective potential leads to an approximate nuclear density. Inspired by discussions with Hardy, we explore the factorization idea for arbitrary many-body Hamiltonians, generalizing the electron-nuclei case, with a focus on the static case. While the exact equations do not lead to any practical advantage, they are illuminating, and may therefore constitute a suitable starting point for approximations. In particular, we find that unitary transformations that diagonalize the coupling term for one of the sub-systems make exact factorization appealing. The algorithms by which the equations for the separate subsystems can be solved in the time-independent case are also explored. We illustrate our discussions using the two-site Holstein model and the quantum Rabi model. Two factorization schemes are possible: one where the boson field feels a potential determined by the electrons, and the reverse exact factorization, where the electrons feel a potential determined by the bosons; both are explored in this work. A comparison with a self-energy approach is also presented.

1 Introduction

The quantum many-body problem [1,2,3] is challenging because of the coupling between all constituents and associated dimensionality curse. Many different approaches have been designed to deal with this problem. Often, they benefit tremendously from the possibility to separate a global system into different subsystems. However, most often such a separation cannot be done exactly, and coupling terms remain, which force some kind of approximation. All embedding approaches, which are used in a large variety of situations such as transport through molecules coupled to leads [4], or the downfolding of electronic bands to a Hubbard model used in dynamical mean field theory [5], make use of the possibility to separate degrees of freedom approximately.

Considering the subsystems on equal footing, products of functions from the independent basis sets can be considered: the coupling term is either treated as a perturbation, or the whole system is expressed in a (truncated) product basis set [6]. Alternatively, the system can be separated with a dissymmetric procedure, based on some physical consideration. In such case, one system is handled under the condition of the other system being in some state. An effective treatment for the former system appears, in which the degrees of freedom of the other system have been integrated out. For example, when constructing pseudopotentials [7] one separates the strongly bound core electrons from the loosely bound valence electrons, which are responsible for chemical bonding. In this approach, usually, the modification

and feedback of the core electrons due to the fact that the valence electrons rearrange when the atoms join to form a cluster or solid is neglected.

Perhaps the most well-known example of the dissymmetric splitting route is the Born-Oppenheimer (BO) approximation, see Ref. [8]: the electronic wavefunction and Hamiltonian depend parametrically on the position of the nuclei, and generate an effective potential for their dynamics. The BO approximation can be turned into an expansion, in which the total wavefunction is decomposed exactly in terms of different electronic states, ground as well as excited, each depending parametrically on the nuclei positions.

In view of retaining exactness as far as possible, still avoiding perturbation theory or expansion over a basis set, Hunter [9] had pointed out, for a general system and Hamiltonian, that a splitting in term of a conditional amplitude (e.g. the electronic wavefunction in case of the BO splitting) and a marginal amplitude (e.g. the nuclei wavefunction in case of the BO splitting) yields an exact effective equation for the marginal amplitude, after integrating the conditional amplitude. However, an equation for the latter was not given. By contrast, focusing on the time-independent electron-nuclei system, Gidopoulos and Gross [10,11] pointed out that, in this case, an effective equation also exists for the conditional amplitude, a result later extended to the time-dependent case in landmark articles [12,13] by Abedi, Maitra and Gross. A Berry gauge field appeared in the separate equations for the subsystems, also generalizing a result from the adiabatic approach [14].

These works have established the 'exact factorization' approach on firm grounds, for the treatment of electron-nuclei non-adiabatic effects, and have been followed by numerous publications. Let us point out, among others, the reverse exact factorization [15] in which the role of electrons and nuclei are exchanged, the many-stage generalization of the conditional-marginal splitting [16], the exact single-electron approach [17,18] and, perhaps the most important achievement, the introduction of a (semi-)classical treatment of the nuclei degrees of freedom, following Bohm [19], thanks to independent trajectories (IT) [20,21] or coupled trajectories (CT) [22,23,24].

In the latter semi-classical treatments, electron-nuclei feedbacks are incorporated, and treated beyond the adiabatic approximation, if not in an exact way. Although most of the applications focused until recently on simplified systems, like the Metiu-Shin model [22], the H_2^+ molecule [12,13], or the $e \otimes E$ Jahn-Teller model [25,26], one year ago the photoinduced ring opening process in oxirane has been studied successfully [24]. In the latter, the focus was on the analysis of decoherence effects, with clear improvement over other existing algorithms such as Tully's 'fewest-switches surface hopping' [27] or 'corrected' versions of it [28] and, of course, over approaches in which decoherence effects are ignored, like Ehrenfest dynamics [29]. Another interesting area of application is the strong-field dynamics of molecules. Indeed, the time-dependent potential energy surface that emerges from the theory provides a strong interpretive tool, as demonstrated in the analysis of the charge resonance enhanced ionization [30,31], laser induced electron-localization [32], and in the justification of the empirical surface hopping method for the laser driven molecular dynamics [33]. For completeness, let us also mention the conditional decomposition approach, as introduced by Albareda and co-workers [34], whose conceptual differences with the exact factorization approach is highlighted in Ref. [34].

Altogether, the 'exact factorization' approach appears hence to be very promising, and it has attracted much attention recently. Still, much remains to be explored. On one side, it is interesting to put the idea of an 'exact' factorization into a more general context, beyond the specific electron-nuclei problem: after all, if no further requirement is specified, the separation of a probability amplitude into a marginal one and a conditional one is trivial and not useful *per se*. On the other side, this is in any case a mere rewriting of the full many-body problem, and only if clever approximations are made, one can expect to gain something on the practical side. The question is then whether the approach does indeed suggest such clever approximations.

Another interesting question is to compare with what is also very often done to deal with the coupling problem, namely, a formulation in terms of a frequency-dependent self-energy or effective interaction [1,2,3,4], which can represent a missing region of space, missing bands, missing excitations, and much more. The frequency-dependence of the self-energy is a crucial ingredient, and when a static self-energy appears in an approach, this is in general a rough approximation. The factorization approach, instead, does not make any additional frequency-dependence appear. How can we situate the frequency-dependent self-energy with respect to the frequency-independent factorization approach? This is a question that Hardy Gross has put

forward a few years ago, and it seems appropriate to include this discussion in a birthday paper.

The present work is intended to contribute to these various discussion points. In order to elucidate general features of the approach, we mainly work with a time-independent hamiltonian, but as we will see, this does not hamper the generality of our discussion, and indeed, we will also touch upon the time-dependent extension. The general discussions directly linked to the exact factorization are contained in Sec. 2. In Sec. 3 we illustrate our points by performing exact and approximate calculations of the electron and boson densities of simple electron-boson coupling Hamiltonians, the two-site Holstein model [35,36], and the quantum Rabi Hamiltonian or two-level pseudo Jahn-Teller model [37,38,39,40,41,42,43]. which leads to further discussions, especially concerning possible approximation schemes. We do not aim at providing breakthrough in the exact factorization area of research, but to mention some relatively less-explored paths. Of course different Hamiltonians may suggest different approximation strategies; however, these examples are meant to give a feeling for the way one might want to deal with the various problems that can appear.

2 General formalism

2.1 Wavefunction and density

Let us consider systems for which two components can readily be identified. This might be because we have objects of two kinds, b and c, that are distinguishable. More generally, for a two-component system, a complete basis for the full wavefunction is formed by products of the separate basis elements for the two components:

$$|N\rangle = \sum_{\mathbf{ij}} C_{\mathbf{ij}} |\mathbf{i}\rangle |\mathbf{j}\rangle, \qquad (1)$$

where the vectors $\mathbf{i}$ and $\mathbf{j}$ label the sets $(i_1, i_2, ...i_{N_i})$ and $(j_1, j_2, ...j_{N_j})$ of quantum numbers that define a basis element in the space of the first and second component, respectively. In the case of fermions, these basis elements can be the Slater determinants built with single-particle states labeled by i_{k_i} and j_{k_j}. In the electron-nuclei case expressed in real space, the basis is composed of delta-functions, and i_k labels the 3 cartesian coordinates of electron k, whereas j_ℓ stands for the 3 cartesian coordinates of nucleus ℓ; in that case, $C_{\mathbf{ij}}$ corresponds to the many-body wavefunction in real space.

We now make an ansatz of factorization of the coefficients C that reflects the idea of conditional and marginal probabilities:

$$C_{\mathbf{ij}} \equiv \pi_{\mathbf{ij}} p_{\mathbf{j}}. \qquad (2)$$

This ansatz is trivially exact if we chose $p = 1$, $\pi = C$. However, in the spirit of the exact factorization here we demand that

$$\sum_{\mathbf{j}} |p_{\mathbf{j}}|^2 = 1 \qquad \text{and} \qquad \sum_{\mathbf{i}} |\pi_{\mathbf{ij}}|^2 = 1 \qquad \text{for all } \mathbf{j}.$$

$$(3)$$

The latter sum, valid for all $\mathbf{j}$ is often referred to as a *partial normalization condition*. Since $\sum_{\mathbf{ij}} |C_{\mathbf{ij}}|^2 = 1$, the conditions Eq. (3) can indeed be fulfilled, e.g. using the choice

$$p_{\mathbf{j}} = e^{i\alpha(\mathbf{j})} \sqrt{\sum_{\mathbf{i}} |C_{\mathbf{ij}}|^2}, \qquad (4)$$

where the phase factor $\alpha(\mathbf{j})$ is a real function. $|p_{\mathbf{j}}|^2$ is the probability to find the system with its second component in state $\mathbf{j}$, irrespective of the value of $\mathbf{i}$, which is a *marginal* probability. $|\pi_{\mathbf{ij}}|^2$ is the *conditional* probability to find the first component of the system in state $\mathbf{i}$, if the second component is in state $\mathbf{j}$. We will refer to the component $p_{\mathbf{j}}$ as being the *marginal amplitude* of the wavefunction, and the component $\pi_{\mathbf{ij}}$ as being the *conditional amplitude* of the wavefunction.

Of course, at variance with such choice of p and π, we might have exchanged the marginal and conditional amplitudes, switching the role of $\mathbf{i}$ and $\mathbf{j}$. This possibility constitutes a first 'variation' on the exact factorization theme.

There is another obvious degree of freedom, associated with the separate choices of basis sets for the $\mathbf{i}$ and $\mathbf{j}$ components. Indeed, the total wavefunction is left invariant under unitary transformations of the basis functions. One can consider invariance with respect to the $\mathbf{i}$ component,

$$
\begin{aligned}
|N\rangle &= \sum_{\mathbf{ij}} C_{\mathbf{ij}} |\mathbf{ij}\rangle \\
&= \sum_{\mathbf{ij}} \Big(\sum_{\tilde{\mathbf{i}}} U_{\mathbf{i}\tilde{\mathbf{i}}}^* C_{\tilde{\mathbf{i}}\mathbf{j}}\Big)\Big(\sum_{\tilde{\mathbf{i}}'} U_{\mathbf{i}\tilde{\mathbf{i}}'} |\tilde{\mathbf{i}}'\mathbf{j}\rangle\Big) = \sum_{\tilde{\mathbf{ij}}} \tilde{C}_{\tilde{\mathbf{i}}\mathbf{j}} |\tilde{\mathbf{i}}\mathbf{j}\rangle,
\end{aligned} \quad (5)
$$

or similarly, with respect to the $\mathbf{j}$ component.

Another choice, a third 'variation' in the exact factorization theme is to consider the static case or the time-dependent case.

All these are moreover subject to the freedom to choose the phase of p_j and π_{ij}, provided the coefficients C are left invariant,

$$C_{\mathbf{ij}} = \pi_{\mathbf{ij}} p_{\mathbf{j}} = (e^{-i\theta_{\mathbf{j}}} \pi_{\mathbf{ij}})(e^{i\theta_{\mathbf{j}}} p_{\mathbf{j}}) = \tilde{\pi}_{\mathbf{ij}} \tilde{p}_{\mathbf{j}}, \qquad (6)$$

where θ is a real-valued function.

In the case of two kinds of distinguishable particles, we denote the creation operator of the first family with $c_i^\dagger$, whereas the creation operator of the second family is $b_j^\dagger$. The many-body basis of the coupled system of $N = N_c + N_b$ particles is the set

$$|\mathbf{ij}\rangle = c_{i_1}^\dagger ... c_{i_{N_c}}^\dagger b_{j_1}^\dagger ..... b_{j_{N_b}}^\dagger |0\rangle, \qquad (7)$$

where the i_s are all different when the $c^\dagger$ create fermions, whereas they can be equal in the case of bosons; the same holds for the $b^\dagger$. The vacuum $|0\rangle = |0\rangle|0\rangle$ is the product of the c- and the b-vacuum. Therefore the many-body wavefunction reads

$$|N\rangle = \sum_{\mathbf{j}} p_{\mathbf{j}} b_{j_1}^\dagger ..... b_{j_{N_b}}^\dagger \sum_{\mathbf{i}} \pi_{\mathbf{ij}} c_{i_1}^\dagger ... c_{i_{N_c}}^\dagger |0\rangle. \qquad (8)$$

Let us now calculate static correlation functions. We start with the usual definition of a density matrix for the species c,

$$n_{k\ell} \equiv \langle N | c_k^\dagger c_\ell | N \rangle. \qquad (9)$$

We now commute the operators such that all b's are on the left and all c's on the right. Since the number of b's and c's is conserved, inserting the identity equals inserting $|0\rangle\langle 0|$ between the two groups of operators. The expectation value of species b integrates out:

$$\langle 0 | b_{j_{N_b}} ..... b_{j_1} b_{j_1'}^\dagger ..... b_{j_{N_b}'}^\dagger |0\rangle = \delta_{\mathbf{j},\mathbf{j}'}, \qquad (10)$$

where permutations in the index vectors are allowed. The result is

$$n_{k\ell} = \sum_{\mathbf{j}} |p_{\mathbf{j}}|^2 \sum_{\mathbf{ii}'} \pi_{\mathbf{ij}}^* \pi_{\mathbf{i}'\mathbf{j}} \langle 0 | c_{i_{N_c}} .... c_{i_1} c_k^\dagger c_\ell c_{i_1'}^\dagger ... c_{i_{N_c}'}^\dagger |0\rangle. \qquad (11)$$

Non-zero contributions are given by $\ell \in \mathbf{i}' \wedge k \in \mathbf{i}$. Moreover, for $\ell = i_0$ and $k = i_0'$ we must have $\mathbf{i} = \mathbf{i}'$ for all other (excluding i_0, i_0') elements of the index vectors. This is the usual expression for the density matrix of species c, except for the fact that in absence of coupling to the species b the coefficients π would not depend on $\mathbf{j}$. The dependence on $\mathbf{j}$ can be translated as "what is the density matrix of c when species b is in the configuration $\mathbf{j}$?".

Up to here, not much seems to be gained: we need both the p and the π in order to calculate the density, in other words, we need the full wavefunction. However, suppose now that we are interested in the one-body density matrix of species b. It reads

$$
\begin{aligned}
n_{k\ell}^b &\equiv \langle N | b_k^\dagger b_\ell | N \rangle \\
&= \sum_{\mathbf{jj}'} p_{\mathbf{j}}^* p_{\mathbf{j}'} \sum_{\mathbf{i}} \pi_{\mathbf{ij}}^* \pi_{\mathbf{ij}'} \langle 0 | b_{j_{N_b}} .... b_{j_1} b_k^\dagger b_\ell b_{j_1'}^\dagger ... b_{j_{N_b}'}^\dagger |0\rangle,
\end{aligned} \quad (12)
$$

since now we have $\delta_{\mathbf{ii}'}$. For $\ell = k$, only $\mathbf{j}' = \mathbf{j}$ contributes. Therefore, the normalization condition for the π can be applied. The final result is

$$n_k^b = N \sum_{j_2, ... j_N} |p_{kj_2 .... j_N}|^2, \qquad (13)$$

i.e., the (generalized) density (or, equivalently, two-body or higher order densities for the species b) for the marginal amplitude can be calculated from the coefficients p alone. This is of course also directly obtained from Eq. (4), since by definition $p_{\mathbf{j}}$ is, besides a phase factor, the square root of the N_b-body density matrix. However, the derivation above is instructive, since it shows why this simple relation is *not* true for the off-diagonal elements of the density *matrix*, or for any other observable, besides functions of the density. It is very important to stress this fact, since it delimits the scope of the exact factorization scheme, at least in its present form.

2.2 Equation for the marginal amplitude

The next step is to consider a time-independent Schrödinger equation, and to proceed to the separation between the marginal and conditional amplitudes. We start from

$$H = {}^cH + {}^bH + {}^{cb}H, \qquad (14)$$

where ^cH acts only on the $\mathbf{i}$ set of orthonormal basis functions, and ^bH acts only on the $\mathbf{j}$ set of orthonormal basis functions. In the distinguishable particle case introduced in the previous subsection, ^cH contains only combinations of operators $c, c^\dagger$, ^bH contains only combinations of operators $b, b^\dagger$, and ^{cb}H contains the mixed terms. Written in the basis Eq. (7), this leads to a Schrödinger equation

$$\sum_{\mathbf{i'j'}} H_{\mathbf{ij}}^{\mathbf{i'j'}} A_{\mathbf{i'j'}}^\nu = E_\nu A_{\mathbf{ij}}^\nu, \tag{15}$$

where we can separate the hamiltonian in the three contributions:

$$H_{\mathbf{ij}}^{\mathbf{i'j'}} = {}^cH_{\mathbf{i}}^{\mathbf{i'}} \delta_{\mathbf{jj'}} + {}^bH_{\mathbf{j}}^{\mathbf{j'}} \delta_{\mathbf{ii'}} + {}^{cb}H_{\mathbf{ij}}^{\mathbf{i'j'}}. \tag{16}$$

Following Hunter [9], we act on the Schrödinger equation with $\sum_{\mathbf{i}} \pi_{\mathbf{ij}}^*$. This yields

$$\sum_{\mathbf{i}} \pi_{\mathbf{ij}}^* \sum_{\mathbf{i'j'}} [{}^cH_{\mathbf{i}}^{\mathbf{i'}} \delta_{\mathbf{j,j'}} + {}^bH_{\mathbf{j}}^{\mathbf{j'}} \delta_{\mathbf{i,i'}} + {}^{cb}H_{\mathbf{ij}}^{\mathbf{i'j'}}]\pi_{\mathbf{i'j'}} p_{\mathbf{j'}}$$
$$= E \sum_{\mathbf{i}} \pi_{\mathbf{ij}}^* \pi_{\mathbf{ij}} p_{\mathbf{j}}, \tag{17}$$

where we have dropped the state index ν. Using the normalization Eq. (3) we obtain

$$\sum_{\mathbf{j'}} \overline{H}_{\mathbf{j}}^{\mathbf{j'}} p_{\mathbf{j'}} = E p_{\mathbf{j}}, \tag{18}$$

where the overline indicates π-weighted contraction with respect to the $\mathbf{i}$ and $\mathbf{i'}$ indices,

$$\overline{H}_{\mathbf{j}}^{\mathbf{j'}} \equiv \sum_{\mathbf{i}} \pi_{\mathbf{ij}}^* \sum_{\mathbf{i'}} \left[{}^cH_{\mathbf{i}}^{\mathbf{i'}} \delta_{\mathbf{jj'}} + {}^bH_{\mathbf{j}}^{\mathbf{j'}} \delta_{\mathbf{ii'}} + {}^{cb}H_{\mathbf{ij}}^{\mathbf{i'j'}} \right] \pi_{\mathbf{i'j'}}. \tag{19}$$

We have hence found, with Eq. (18), an effective Schrödinger equation for the component $\mathbf{j}$ alone, where $\mathbf{i}$ is completely integrated out. It contains three terms, the first of which is diagonal in $\mathbf{j}$,

$$^c\epsilon_{\mathbf{j}} \equiv \sum_{\mathbf{ii'}} \pi_{\mathbf{ij}}^* {}^cH_{\mathbf{i}}^{\mathbf{i'}} \pi_{\mathbf{i'j}}, \tag{20}$$

while the two others terms couple the $\mathbf{j}$ states,

$$^b\overline{H}_{\mathbf{j}}^{\mathbf{j'}} \equiv {}^bH_{\mathbf{j}}^{\mathbf{j'}} \sum_{\mathbf{i}} \pi_{\mathbf{ij}}^* \pi_{\mathbf{ij'}}, \tag{21}$$

and

$$^{cb}\overline{H}_{\mathbf{j}}^{\mathbf{j'}} \equiv \sum_{\mathbf{ii'}} \pi_{\mathbf{ij}}^* {}^{cb}H_{\mathbf{ij}}^{\mathbf{i'j'}} \pi_{\mathbf{i'j'}}. \tag{22}$$

The diagonal elements of both $^b\overline{H}$ and $^{cb}\overline{H}$, that are invariant with respect to phase changes (see Appendix A), can be combined with $^c\epsilon_{\mathbf{j}}$ to define an overall diagonal term

$$\epsilon_{\mathbf{j}} \equiv {}^c\epsilon_{\mathbf{j}} + {}^b\bar{H}_{\mathbf{j}}^{\mathbf{j}} + {}^{cb}\bar{H}_{\mathbf{j}}^{\mathbf{j}}. \tag{23}$$

The final effective Schrödinger equation for the marginal amplitude $p_{\mathbf{j}}$ writes

$$\epsilon_{\mathbf{j}} p_{\mathbf{j}} + \sum_{\mathbf{j'}\neq\mathbf{j}} ({}^b\overline{H}_{\mathbf{j}}^{\mathbf{j'}} + {}^{cb}\overline{H}_{\mathbf{j}}^{\mathbf{j'}}) p_{\mathbf{j'}} = E p_{\mathbf{j}}. \tag{24}$$

This equation is phase-contravariant, thanks to the combined use of Eqs. (131), (132) and (133), while the diagonal term and eigenenergy are phase-invariant.

An important class of hamiltonians is given by the case where the cb-coupling is potential-like, namely $^{cb}H_{\mathbf{ij}}^{\mathbf{i'j'}} = {}^{cb}H_{\mathbf{ij}}\delta_{\mathbf{ii'}}\delta_{\mathbf{jj'}}$. In that case, the coupling between the two systems merely adds the potential energy surface $\epsilon_{\mathbf{j}}$ with respect to the π-weighted $^b\overline{H}$:

$$\epsilon_{\mathbf{j}} p_{\mathbf{j}} + \sum_{\mathbf{j'}\neq\mathbf{j}} {}^b\overline{H}_{\mathbf{j}}^{\mathbf{j'}} p_{\mathbf{j'}} = E p_{\mathbf{j}}. \tag{25}$$

Actually, this also holds if ^{cb}H is diagonal only in $\mathbf{j}$, but not in $\mathbf{i}$. This will be further elaborated in Sec. 2.6.

2.3 Equation for the conditional amplitude

Beyond the work of Hunter [9], who wrote the equation corresponding to Eq. (18), two routes are possible in order to obtain an equation for π. The simplest one comes from the division of Eq. (15) by $p_{\mathbf{j}}$,

$$\sum_{\mathbf{i'j'}} \left(\frac{1}{p_{\mathbf{j}}} H_{\mathbf{ij}}^{\mathbf{i'j'}} p_{\mathbf{j'}}\right) \pi_{\mathbf{i'j'}} = E\pi_{\mathbf{ij}}, \tag{26}$$

which might not be always possible, however. We will denote generically by the letter 'h', matrix elements of H or its components pre-divided by $p_{\mathbf{j}}$ and post-multiplied by $p_{\mathbf{j'}}$, giving

$$\sum_{\mathbf{i'j'}} h_{\mathbf{ij}}^{\mathbf{i'j'}} \pi_{\mathbf{i'j'}} = E\pi_{\mathbf{ij}}. \tag{27}$$

Such an equation is however as complex as the original one, Eq. (15), as the conditional amplitudes π are as numerous and as tightly coupled as the original C. The Hamiltonian h is even not hermitian anymore. Still, now, different approaches and adequate approximations could be considered separately for Eqs. (24) and (27), allowing new strategies. However, it is not obvious how to handle the coupled Eqs. (24) and (27) for p and π, with the same total energy E appearing as eigenvalue, especially when different approximations are done on the two equations.

This latter problem can be solved easily. Indeed, E can be eliminated from Eq. (27) by using Eq. 18 divided by $p_{\mathbf{j}}$, then Eq. (23), giving

$$\sum_{\mathbf{i'j'}} \left(h_{\mathbf{ij}}^{\mathbf{i'j'}} - \bar{h}_{\mathbf{j}}\delta_{\mathbf{ii'}}\delta_{\mathbf{jj'}}\right)\pi_{\mathbf{i'j'}} = \epsilon_{\mathbf{j}}\pi_{\mathbf{ij}}. \tag{28}$$

Generically, $\bar{h}$ will denote $\bar{H}$ matrix elements pre-divided by $p_{\mathbf{j}}$, post-multiplied by $p_{\mathbf{j'}}$, and summed over $\mathbf{j'}$ except $\mathbf{j}$:

$$\bar{h}_{\mathbf{j}} \equiv \frac{1}{p_{\mathbf{j}}} \sum_{\mathbf{j'}\neq\mathbf{j}} \overline{H}_{\mathbf{j}}^{\mathbf{j'}} p_{\mathbf{j'}}. \tag{29}$$

Note that $\bar{h}_{\mathbf{j}}$ is phase-invariant.

Before detailing the corresponding equations in terms of ^cH, ^cH, and ^{cb}H, let us explain the second route. We consider the well-known stationarity property for the expectation value of H with respect to constrained variations of the wavefunctions

$C_{\mathbf{ij}}$. Following Gidopoulos and Gross [10,11], when the decomposition Eq. (2) is considered, $\mathbf{j}$-specific Lagrange parameters $\lambda_{\mathbf{j}}$ appear, each of these corresponding to the $\mathbf{j}$-associated normalization constraint for $\pi_{\mathbf{ij}}$, Eq. (3), as well as a global Lagrange parameter Λ associated with the marginal amplitudes,

$$
E^L[\pi_{\mathbf{ij}}, p_{\mathbf{j}}, \lambda_{\mathbf{j}}, \Lambda] = \Big(\sum_{\mathbf{ij}} \sum_{\mathbf{i'j'}} \pi_{\mathbf{ij}}^* p_{\mathbf{j}}^* H_{\mathbf{ij}}^{\mathbf{i'j'}} \pi_{\mathbf{i'j'}} p_{\mathbf{j'}}
$$
$$
- \sum_{\mathbf{j}} \lambda_{\mathbf{j}} \Big(\sum_{\mathbf{i}} |\pi_{\mathbf{ij}}|^2 - 1 \Big)
$$
$$
- \Lambda \Big(\sum_{\mathbf{j}} |p_{\mathbf{j}}|^2 - 1 \Big) \Big), \tag{30}
$$

where E^L denotes the Hamiltonian expectation functional, augmented by the Lagrange terms. When $\pi_{\mathbf{ij}}$ and $\lambda_{\mathbf{j}}$ are solutions of Eq. (15) and fulfill the adequate normalization conditions, then

$$
\delta E^L[\pi_{\mathbf{ij}}, p_{\mathbf{j}}, \lambda_{\mathbf{j}}, \Lambda] = 0. \tag{31}
$$

Variation of Eq. (30) with respect to $p_{\mathbf{j}}^*$ and $\pi_{\mathbf{ij}}^*$ yields

$$
(\underline{\nabla}_p^L)_{\mathbf{j}} = \frac{\partial E^L}{\partial p_{\mathbf{j}}^*} = \sum_{\mathbf{j'}} \overline{H}_{\mathbf{j}}^{\mathbf{j'}} p_{\mathbf{j'}} - \Lambda p_{\mathbf{j}} \tag{32}
$$

$$
(\underline{\nabla}_{\pi\mathbf{j}}^L)_{\mathbf{i}} = \frac{\partial E^L}{\partial \pi_{\mathbf{ij}}^*} = \Big(p_{\mathbf{j}}^* \sum_{\mathbf{i'j'}} H_{\mathbf{ij}}^{\mathbf{i'j'}} p_{\mathbf{j'}} \Big) \pi_{\mathbf{i'j'}} - \lambda_{\mathbf{j}} \pi_{\mathbf{ij}}. \tag{33}
$$

Requiring Eq. (32) to vanish, following Eq. (31), yields Eq. (18). Requiring Eq. (33) to vanish, then dividing it by $|p_{\mathbf{j}}|^2$ -which might not always be possible-, yields an equation for $\pi_{\mathbf{ij}}$

$$
\sum_{\mathbf{i'j'}} h_{\mathbf{ij}}^{\mathbf{i'j'}} \pi_{\mathbf{i'j'}} = \lambda_{\mathbf{j}}' \pi_{\mathbf{ij}}, \tag{34}
$$

with $\lambda_{\mathbf{j}}' = \lambda_{\mathbf{j}}/|p_{\mathbf{j}}|^2$. Apparently, this equation differs from Eq. (27). However, pre-multiplication by $\pi_{\mathbf{ij}}^*$ and elimination of $\lambda_{\mathbf{j'}}$ yields the same Eq. (28), that can be written more explicitly in terms of its components,

$$
\sum_{\mathbf{i'}} {}^c H_{\mathbf{i}}^{\mathbf{i'}} \pi_{\mathbf{i'j}} + \Big(\sum_{\mathbf{j'}} {}^b h_{\mathbf{j}}^{\mathbf{j'}} \pi_{\mathbf{ij'}} - {}^b \overline{h}_{\mathbf{j}} \pi_{\mathbf{ij}} \Big)
$$
$$
+ \Big(\sum_{\mathbf{i'j'}} {}^{cb} h_{\mathbf{ij}}^{\mathbf{i'j'}} \pi_{\mathbf{i'j'}} - {}^{cb} \overline{h}_{\mathbf{j}} \pi_{\mathbf{ij}} \Big) = \epsilon_{\mathbf{j}} \pi_{\mathbf{ij}}. \tag{35}
$$

Globally, Eq. (35) is phase-covariant. Since the outcome of the two routes is the same, one might wonder why to bother about the variational approach. Actually, the latter provides an insight on how to actually solve the coupled marginal and conditional equations, as we shall see later.

Eq. 35 can be further simplified by considering specificities of the Hamiltonian, like the above-mentioned diagonal character of the ${}^{cb}H$ component of the Hamiltonian, or the smallness of the nuclei kinetic energy, due to the large nuclei-electron mass ratio. Before detailing such simplifications, let us analyze some global properties of the coupled equations Eqs. (24) and (35), and discuss what might have been gained at this global level.

2.4 Coupling of the marginal and conditional equations

In the general case, there is a feedback from the marginal equation to the conditional equation, and vice-versa, as the matrix elements found in one equation have been built from the knowledge of the wavefunction from the other equation. Of course, there are specific cases where the separation is perfect, see an example in Sec. 3.6, or where the specificities of the Hamiltonian allows one, as a reasonable approximation, to discard the coupling, or to treat it as a perturbation on the basis of the separated entities.

Let us focus, however, on the exact factorization feedback process, and the algorithm by which one can address the joint handling of Eq. (24) and Eq. (35), each of these being possibly approximated, albeit *not* suppressing the feedback loop. In the exact factorization spirit, one might consider solving them sequentially, by an iterative process. Updating them in parallel should be considered as well. These two possibilities constitute new variations of the theme.

As a starting point of a sequential iterative approach, one might consider an ansatz for the $p_{\mathbf{j}}$, to be used to build the matrix elements of Eq. (35). However, none of them can be zero as already outlined. Even if none of them vanishes, perhaps having some of them small might be problematic, as some h elements might become unphysically large to start an iterative procedure.

The alternative route is more appealing. It consists in considering an ansatz for $\pi_{\mathbf{ij}}$, that fulfills the partial normalization conditions Eq. (3), building the matrix elements needed in Eq. (24), solving it, and then coming back to Eq. (35).

Still, the division by $p_{\mathbf{j}}$ needed to obtain the matrix elements in Eq. (35) might not be eliminated. Considering first a discrete $\mathbf{j}$-set, some of the $p_{\mathbf{j}}$ of the final solution might be exactly zero, either accidentally or due to some symmetry properties of the Hamiltonian. In the latter case, the set of vanishing $p_{\mathbf{j}}$ might possibly be determined a priori, and actually discarded from the admitted set of $\mathbf{j}$, because all corresponding $C_{\mathbf{ij}}$ vanish. We will encounter such a case in Sec. 3.7. If the vanishing is accidental, we remark anyhow that $C_{\mathbf{ij}} = 0$ as well, which means the sum in Eq. 26 also vanishes. All the corresponding $\pi_{\mathbf{ij}}$ are undetermined, but do not diverge due to their normalization. This might possibly yield convergence problems, or difficulties with approximations that would not preserve the indeterminacy. It is a potential problem, to be treated on a case-by-case basis. The problem is even more acute in the case of a continuous $\mathbf{j}$-set, as the marginal amplitudes cannot change sign without vanishing (in the real case). At the sign change, the off-diagonal elements diverge. This problem goes away if $p_{\mathbf{j}}$ never vanishes.

If one now supposes that the set of $p_{\mathbf{j}}$ does not vanish, neither at initialization nor during or at convergence, one will face the challenge to solve Eq. (35), at fixed values of $p_{\mathbf{j}}$: a set of coupled equations, each $\mathbf{j}$-equation being possibly considered as stand-alone, with the two terms in parentheses giving the influence of $\mathbf{j'}$ on the $\mathbf{j}$ quantities. One might consider thus solving one $\mathbf{j}$-equation under the condition of fixing the π for all $\mathbf{j'} \neq \mathbf{j}$. The set of π-equations could be either addressed sequentially or in parallel. Numerical experimentations, to be described in Sec. 3.8, have alerted us on the inherent instability of such a fixed $\mathbf{j'}$ approach to the $\mathbf{j}$-equation.

Indeed, with the appearance of ${}^{b}h_{\mathbf{j}}^{\mathbf{j}'}$ and ${}^{bc}h_{\mathbf{ij}}^{\mathbf{i}'\mathbf{j}'}$, the equation (35), which is to be solved for each $\mathbf{j}$ separately is not any longer a usual Schrödinger equation. Fixing the $\mathbf{j}'$ terms in the left-hand side gives an equation of the type

$$\underline{\underline{A}}x + \underline{y} = \epsilon \underline{x} \qquad (36)$$

under constraint $\underline{x}^{\dagger}\underline{x} = 1$, whose peculiar mathematical structure is analyzed in Appendix B. As explained in Sec. 2.7, its time-dependent generalization has the mathematical structure of a time-dependent Schrödinger equation with an inhomogeneous term. Remaining with the time-independent Eq. (36), Appendix B shows that the effect of a change of $\underline{y}$ can be amplified largely, in the case ϵ is close to one of the eigenvalues of $\underline{\underline{A}}$. Thus, addressing such j-coupled equations in an iterative fashion (be it sequential or parallel), in which the solution of one equation at one specific step will be fed into the other equations for the next step, without specific treatment (e.g. without damping) might naturally lead to difficulties in convergence (see e.g. Ref. [44]).

Instead of solving for $\pi_{\mathbf{ij}}$ with fixed $\mathbf{j}'$ conditions, one can update all the π jointly, under the guidance of the variational principle Eq. (31). Such procedure is inherently more stable. In addition, a variational principle gives not only equations to be fulfilled at the minimum, but also gradients to be followed when one is not yet at convergence, thus providing an iterative flow, similar to the time-dependent case of Sec. 2.7. Among the optimization algorithms that have been designed to address such non-linear optimization problems, the preconditioned conjugate-gradient algorithm enjoys stability, unconditional convergence, optimality in case of purely quadratic problems, with affordable overhead linked to accurate line minimization for non-linear problems, see Refs. [45] and chapter 10 of [46]. We will use such algorithm in Sec. 3.8. In view of this application, we derive now the so-called preconditioned gradients $\underline{G}^{L}$ of the Lagrange functional E^{L} Eq. (30), generalizing the results of Refs. [47] and [48]. Preconditioning is the application of a transformation of the search space, that has the goal to reduce the condition number of the problem, i.e. to make the search direction from the starting state closer to the direction of the minimum[48,44].

Concerning the marginal amplitudes, we define $\underline{\underline{P}}_{\perp p}$, the projector on the subspace orthogonal to p,

$$\left(\underline{\underline{P}}_{\perp p}\right)_{\mathbf{j}}^{\mathbf{j}'} = \delta_{\mathbf{jj}'} - p_{\mathbf{j}}p_{\mathbf{j}'}^{*}. \qquad (37)$$

We define also the preconditioner $\underline{\overline{K}}_{p}^{-1}$, where the $\underline{\overline{K}}_{p}$ operator is an approximation of $\underline{\underline{P}}_{\perp p}\underline{\overline{H}}\,\underline{\underline{P}}_{\perp p}$. Since the dimensionality of the search space can be huge, such operator should be numerically easy to invert, which is usually obtained by taking a diagonal or quasi-diagonal approximation. In Sec. 3.8 we will use the diagonal approximation. The preconditioned gradient for p, orthogonalized with respect to p, writes

$$\underline{G}_{p} = \underline{\underline{P}}_{\perp p}\underline{\overline{K}}_{p}^{-1}\nabla_{p}^{L}. \qquad (38)$$

In the present notations, $\nabla_{p}^{L} = \underline{\underline{P}}_{\perp p}\underline{\overline{H}}p$. Line minimization is performed in the one-dimensional space obtained by mixing the initial state with this preconditioned gradient direction, followed by normalization of the trial new p states. Note that the normalization is preserved at first order thanks to Eq. 37.

Similarly, we introduce $\underline{\underline{P}}_{\perp \pi \mathbf{j}}$, the projector on the subspace orthogonal to $\pi_{\mathbf{j}}$ inside the $\mathbf{i}$-space,

$$\left(\underline{\underline{P}}_{\perp \pi \mathbf{j}}\right)_{\mathbf{i}}^{\mathbf{i}'} = \delta_{\mathbf{ii}'} - \pi_{\mathbf{ij}}\pi_{\mathbf{i}'\mathbf{j}}^{*}, \qquad (39)$$

as well as a $\mathbf{j}$-diagonal preconditioner $\underline{\overline{K}}_{\pi \mathbf{j}}^{-1}$, where $\underline{\overline{K}}_{\pi \mathbf{j}}$ is an approximation of $\underline{\underline{P}}_{\perp \pi \mathbf{j}}\underline{h}_{\mathbf{j}}^{\mathbf{j}}\underline{\underline{P}}_{\perp \pi \mathbf{j}}$ that is easy to invert. We will use the diagonal elements of $\underline{h}_{\mathbf{j}}^{\mathbf{j}}$ in Sec. 3.8, with an additional global shift to guarantee the strict positive definiteness of this matrix, thus avoiding a potential problem with matrix inversion. The preconditioned gradient for $\pi_{\mathbf{ij}}$ writes

$$\underline{G}_{\pi \mathbf{j}} = \underline{\underline{P}}_{\perp \pi \mathbf{j}}\underline{\overline{K}}_{\pi \mathbf{j}}^{-1}\nabla_{\pi \mathbf{j}}^{L}. \qquad (40)$$

Note that contributions from $\pi_{\mathbf{ij}'}$ are needed to compute $\nabla_{\pi \mathbf{j}}^{L}$.

Having defined the preconditioned gradients, they can be used in different flows. A first possibility is to find the stationary points of Eq. (30) by simultaneous line optimization over all $p_{\mathbf{j}}$ and $\pi_{\mathbf{ij}}$, concurrently, using e.g. the steepest descent algorithm or the conjugate gradient algorithm [47,48]. Another possibility, that we will follow in Sec. 3.8, is to optimize all $p_{\mathbf{j}}$ at fixed $\pi_{\mathbf{ij}}$, then optimize all $\pi_{\mathbf{ij}}$ at fixed $p_{\mathbf{j}}$, and then iterate.

As a final remark on the coupling between marginal and conditional equations, note that the diagonal term $\epsilon_{\mathbf{j}}$ appearing in Eq. (23) is not uniquely defined, as a redefinition

$$\epsilon_{\mathbf{j}} \rightarrow \epsilon_{\mathbf{j}}' = \epsilon_{\mathbf{j}} + \Delta_{\mathbf{j}} \qquad (41)$$

is possible, by which quantities that are multiple of the unit operator in the $\mathbf{i}$ space and diagonal in the $\mathbf{j}$ space are transfered between Eqs. (24) and (35).

2.5 Did we gain something ?

Despite its phase-dependence, Eq. (18) (or Eq. (24) that governs the marginal amplitude, is appealing, since the complexity of the original problem, with the two components $\mathbf{i}$ and $\mathbf{j}$, has been reduced entirely to a problem for the component $\mathbf{j}$ only, provided one can compute the matrix elements ${}^{b}\overline{H}_{\mathbf{j}}^{\mathbf{j}'}$ and ${}^{cb}\overline{H}_{\mathbf{j}}^{\mathbf{j}'}$.

By contrast, Eq. (35) is as complicated as the original Eq. (15), because, in the general case, the cb term, that contains ${}^{cb}h_{\mathbf{ij}}^{\mathbf{i}'\mathbf{j}'}$, couples all pairs of $\pi_{\mathbf{ij}}$, linking $\mathbf{i}$ and $\mathbf{i}'$ as well as $\mathbf{j}$ and $\mathbf{j}'$. The b term, that contains ${}^{b}h_{\mathbf{j}}^{\mathbf{j}'}$ is potentially less harmful, since it does not couple pairs of $\pi_{\mathbf{ij}}$ with differing i components. Still, no direct reduction of the workload is achieved, even if a clever numerical approach is used.

Of course, we cannot expect to have a breakthrough by merely rewriting the many-body problem in an exact way. We can only expect to have a better starting point for approximations, by highlighting some important physics. One of the points that might be important is the following.

The scope of the exact factorization approach is to calculate one- (or many-) body densities of a given species. For two interacting species placed in the empty space, translational invariance renders the one-body density for one species absolutely meaningless: the probability to find a particle in empty space is the same everywhere. What is instead meaningful is the probability to find a particle in a position once the position of the other particles are fixed, or the density of a given species supposing a certain density of the other species, if the two species interact. For example, when calculating the electron density of a crystal usually the positions of the nuclei are decided. This corresponds to a legitimate symmetry breaking, but such a trick is not always possible. This shows that the good question is not "what is the density of the electrons", but rather "what is the density of electrons, given that we have a certain distribution of nuclei?" Such a question is answered when one calculates an at least two-body correlation function. As we have seen, the exact factorization approach is another way to express this "conditional probability" picture very naturally. This means in particular that one might find appropriate iteration schemes that include symmetry breaking "on the fly", and therefore allow one to use approximations that treat correlation on a lower level.

2.6 Semi-diagonal coupling

An exact rewriting of the general many-body problem without further simplifications or at least interpretation that might suggest approximations is not of any practical use. Often a given interpretation is particularly meaningful in some special cases. For this reason, in this subsection we will focus on cb Hamiltonians that have a particular property, which we term 'semi-diagonal coupling', and use the freedom to make a unitary transformations in the space of one component, as expressed in Eq. (5). Eqs. (24) and (35) simplify to a large extent, and decouple partly.

Indeed, suppose that the bc Hamiltonian is such that one can choose a basis set for $\mathbf{j}$ that diagonalizes $^{cb}H_{\mathbf{ij}}^{\mathbf{i'j'}}$ in the $\mathbf{j}$-space, *irrespective* of $\mathbf{i}$ and $\mathbf{i'}$, i.e. $^{cb}H_{\mathbf{i}}^{\mathbf{i'}}(\mathbf{j})$. Then, Eq. (35) reduces to

$$\sum_{\mathbf{i'}} \left({}^{c}H_{\mathbf{i}}^{\mathbf{i'}} + {}^{cb}H_{\mathbf{i}}^{\mathbf{i'}}(\mathbf{j}) + \delta_{\mathbf{ii'}} {}^{b}H_{\mathbf{j}}^{\mathbf{j}} \right)\pi_{\mathbf{i'j}}$$
$$+ \left(\sum_{\mathbf{j'}\neq\mathbf{j}} {}^{b}h_{\mathbf{j}}^{\mathbf{j'}} \pi_{\mathbf{ij'}} \right) = \epsilon_{\mathbf{j}}\pi_{\mathbf{ij}} \tag{42}$$

with

$$\epsilon_{\mathbf{j}} = {}^{c}\epsilon_{\mathbf{j}} + {}^{b}\bar{H}_{\mathbf{j}}^{\mathbf{j}} + {}^{cb}\bar{H}_{\mathbf{j}}^{\mathbf{j}} + {}^{b}\bar{h}_{\mathbf{j}}, \tag{43}$$

where we have used the freedom Eq. (41) to make Eq. (42) as simple as possible.

In Eq. (42), if the non-diagonal $^{b}h_{\mathbf{j}}^{\mathbf{j'}}$ term could be ignored, one could find for each value of $\mathbf{j}$, the energy $\epsilon_{\mathbf{j}}$ by diagonalization with respect to component $\mathbf{i}$ only.

The equation for the marginal amplitude simplifies even more, since there is no off-diagonal $^{cb}\bar{H}$ contribution:

$$\epsilon_{\mathbf{j}}p_{\mathbf{j}} + \sum_{\mathbf{j'}\neq\mathbf{j}} ({}^{b}\bar{H}_{\mathbf{j}}^{\mathbf{j'}})p_{\mathbf{j'}} = Ep_{\mathbf{j}}. \tag{44}$$

The semi-diagonal property of the cb coupling induces a simplification of both equations. Note that the term $\mathbf{j'} = \mathbf{j}$ is discarded from the sum in Eq. (44), as $^{b}\bar{H}_{\mathbf{j}}^{\mathbf{j}} = {}^{b}\bar{h}_{\mathbf{j}}$.

In Eq. (44), the quantity $\epsilon_{\mathbf{j}}$ is a diagonal operator for the marginal amplitude. Moreover, it is phase-invariant. It might be considered as the natural local potential of the system. However a potential operator is defined with respect to a complementary kinetic operator, while the second term in the left-hand side is, at this stage, defined as being the off-diagonal part stemming from the $^{b}\bar{H}$ operator, irrespective of whether or not it is a natural kinetic operator. Actually, any discretization of a kinetic operator in real space has a diagonal contribution, so this identification of $\epsilon_{\mathbf{j}}$ as a local potential has to be understood in a very general way.

A further refinement (or 'variation') of this semi-diagonal scheme is obtained by separating operators contributing to ^{b}H in two sets, which are either diagonal in $\mathbf{j}$ (b-diagonal, or bd), or non-diagonal in $\mathbf{j}$ (b-non-diagonal, or bn). This goes hand in hand with the freedom Eq. (41). We write

$$^{b}H_{\mathbf{j}}^{\mathbf{j'}} = {}^{bd}H(\mathbf{j})\delta_{\mathbf{jj'}} + {}^{bn}H_{\mathbf{j}}^{\mathbf{j'}}. \tag{45}$$

The equation for the conditional amplitude is then

$$\sum_{\mathbf{i'}} \left({}^{c}H_{\mathbf{i}}^{\mathbf{i'}} + {}^{cb}H_{\mathbf{i}}^{\mathbf{i'}}(\mathbf{j}) + \delta_{\mathbf{ii'}} {}^{bd}H(\mathbf{j}) \right)\pi_{\mathbf{i'j}}$$
$$+ \left(\sum_{\mathbf{j'}} {}^{bn}h_{\mathbf{j}}^{\mathbf{j'}} \pi_{\mathbf{ij'}} \right) = \epsilon_{\mathbf{j}}\pi_{\mathbf{ij}}. \tag{46}$$

in which $\epsilon_{\mathbf{j}}$ has been redefined by taking into account Eq. (45).

A further gain is also present if the b Hamiltonian has a particular form, like a derivative operator, and/or if it couples the $\mathbf{j}$ states only to a limited extent, and/or if it is weak with respect to the effect of the $\mathbf{j}$-diagonal part of Eq. (42). Note that the Coulomb potential in real space is a particular case, where $^{cb}H_{\mathbf{ij}}^{\mathbf{i'j'}} = \delta_{\mathbf{ii'}}\delta_{\mathbf{jj'}}$.

2.7 Time-dependent generalization

In Ref. [12] and subsequent publications, the exact factorization has been generalized to the time-dependent case. Indeed, the emphasis has often been put on the time-evolution capability of this formalism.

In the general approach that we outline, Eqs. (35) and (24) transform to the following equations in the time-dependent case :

$$\sum_{\mathbf{i'}} \left({}^{c}H_{\mathbf{i}}^{\mathbf{i'}}(t)\pi_{\mathbf{i'j}}(t) - \epsilon_{\mathbf{j}}(t)\pi_{\mathbf{ij}}(t) \right)$$
$$+ \left(\sum_{\mathbf{j'}} {}^{b}h_{\mathbf{j}}^{\mathbf{j'}}(t)\pi_{\mathbf{ij'}}(t) - {}^{b}\bar{h}_{\mathbf{j}}(t)\pi_{\mathbf{ij}}(t) \right)$$
$$+ \left(\sum_{\mathbf{i'j'}} {}^{cb}h_{\mathbf{ij}}^{\mathbf{i'j'}}(t)\pi_{\mathbf{i'j'}}(t) - {}^{cb}\bar{h}_{\mathbf{j}}(t)\pi_{\mathbf{ij}}(t) \right)$$
$$= i\frac{\partial\pi_{\mathbf{ij}}}{\partial t}, \tag{47}$$

where we have defined the time-dependent generalization of Eq.(43)

$$\epsilon_{\mathbf{j}}(t) = {}^c\epsilon_{\mathbf{j}}(t) + {}^b\overline{H}_{\mathbf{j}}^{\mathbf{j}}(t) + {}^{cb}\overline{H}_{\mathbf{j}}^{\mathbf{j}}(t) + {}^t\epsilon_{\mathbf{j}}(t), \qquad (48)$$

with

$$ {}^t\epsilon_{\mathbf{j}}(t) = \sum_{\mathbf{i}} \pi_{\mathbf{ij}}^*(-i\frac{\partial \pi_{\mathbf{ij}}}{\partial t}), \qquad (49)$$

and

$$\epsilon_{\mathbf{j}}(t)p_{\mathbf{j}}(t) + \sum_{\mathbf{j}'\neq\mathbf{j}} ({}^b\overline{H}_{\mathbf{j}}^{\mathbf{j}'}(t) + {}^{cb}\overline{H}_{\mathbf{j}}^{\mathbf{j}'}(t))p_{\mathbf{j}'}(t) = i\frac{\partial p_{\mathbf{j}}}{\partial t}. \quad (50)$$

For sake of clarity, henceforth, we will not continue to write explicitly the time-dependence of the H matrix elements, but will keep it for the other items.

Following the same line of thought as for the time-independent case, the equations simplify for a semi-diagonal coupling. The equation for the conditional amplitude then becomes

$$\sum_{\mathbf{i}'} \left({}^cH_{\mathbf{i}}^{\mathbf{i}'} + {}^{cb}H_{\mathbf{i}}^{\mathbf{i}'}(\mathbf{j}) + \delta_{\mathbf{ii}'}({}^bH_{\mathbf{j}}^{\mathbf{j}} - \epsilon_{\mathbf{j}}(t)) \right) \pi_{\mathbf{i}'\mathbf{j}}(t)$$
$$+ \left(\sum_{\mathbf{j}'\neq\mathbf{j}} {}^bh_{\mathbf{j}}^{\mathbf{j}'}(t)\pi_{\mathbf{ij}'}(t) \right) = i\frac{\partial \pi_{\mathbf{ij}}}{\partial t}, \qquad (51)$$

with

$$\epsilon_{\mathbf{j}}(t) = {}^c\epsilon_{\mathbf{j}}(t) + {}^{cb}\bar{H}(\mathbf{j},t) + {}^b\bar{H}_{\mathbf{j}}^{\mathbf{j}}(t) + {}^b\overline{h}_{\mathbf{j}}(t) + {}^t\epsilon_{\mathbf{j}}(t), \quad (52)$$

while the equation for the marginal amplitude is also simplified,

$$\epsilon_{\mathbf{j}}(t)p_{\mathbf{j}}(t) + \sum_{\mathbf{j}'} ({}^b\overline{H}_{\mathbf{j}}^{\mathbf{j}'}(t) - \delta_{\mathbf{jj}'}{}^b\overline{h}_{\mathbf{j}}(t))p_{\mathbf{j}'}(t) = i\frac{\partial p_{\mathbf{j}}}{\partial t}. \quad (53)$$

Like in the time-independent case, the diagonal/non-diagonal splitting of ${}^bH$ can be used.

The time-dependent equation for $p_{\mathbf{j}}$ has the mathematical structure of a usual time-dependent Schrödinger equation. At variance, considering one $\mathbf{j}$ at a time, the equation for $\pi_{\mathbf{ij}}$ is a time-dependent Schrödinger equation with an additional source term. Anyhow, solving either equations resorts to ordinary time-dependent numerical techniques.

In both cases, conservation of the norm or the partial norm is obvious, as one deduces $\Re \sum_{\mathbf{i}} \pi_{\mathbf{ij}}^* \frac{\partial \pi_{\mathbf{ij}}}{\partial t} = 0$ from Eq. (51) and Eq. (52) and $\Re \sum_{\mathbf{j}} p_{\mathbf{j}}^* \frac{\partial p_{\mathbf{j}}}{\partial t} = 0$ from Eq. (53).

2.8 Treatment of the electron-nuclei system in real space

We now establish the final link with the works on the exact factorization from Hardy and co-workers, by considering the electron-nuclei interaction in real space. This link is interesting because the electron-nuclei case is at present intensively investigated, and placing it in the present general framework shows what particular properties of the hamiltonian play a role in the electron-nuclei case.

We use now the same notations as in Ref. [12]: electronic real space coordinates $\underline{\mathbf{r}}$ correspond to $\mathbf{i}$ indices, nuclear real space coordinates $\underline{\underline{\mathbf{R}}}$ correspond to $\mathbf{j}$ indices, the electron wavefunction $\Phi_{\underline{\underline{\mathbf{R}}}}(\underline{\mathbf{r}},t)$ in Ref. [12] corresponds to the conditional amplitude $\pi_{\mathbf{ij}}(t)$, and the nuclear wavefunction $\chi(\underline{\underline{\mathbf{R}}},t)$ corresponds to the marginal amplitude $p_{\mathbf{j}}(t)$.

In this context, the coupling Hamiltonian $V_{e-n}(\underline{\mathbf{r}},\underline{\underline{\mathbf{R}}})$ is diagonal in both $\underline{\mathbf{r}}$ and $\underline{\underline{\mathbf{R}}}$, which allows one to take, at will, the electronic wavefunction [12] or the nuclei wavefunction [15] as the conditional amplitude, and the other as the marginal amplitude. Focusing on the first case, the link between our Eqs. (51-53) and Eqs.(6-7) of Ref. [12] can be established by a redefinition of $\epsilon_{\mathbf{j}}$ and by invoking the following equality, relative to the derivative with respect to the nucleus position:

$$\langle \phi_{\underline{\underline{\mathbf{R}}}}(t)| - \partial_\alpha^2|\phi_{\underline{\underline{\mathbf{R}}}}(t)\rangle_{\underline{\mathbf{r}}} \; \chi(\underline{\underline{\mathbf{R}}},t) = $$
$$(-i\partial_\alpha + A_\alpha(\underline{\underline{\mathbf{R}}},t))^2 \chi(\underline{\underline{\mathbf{R}}},t) + $$
$$\chi(\underline{\underline{\mathbf{R}}},t) \; \langle \phi_{\underline{\underline{\mathbf{R}}}}(t)|(-i\partial_\alpha - A_\alpha(\underline{\underline{\mathbf{R}}},t))^2|\phi_{\underline{\underline{\mathbf{R}}}}(t)\rangle_{\underline{\mathbf{r}}}, $$
$$(54)$$

where

$$A_\alpha(\underline{\underline{\mathbf{R}}},t) = \langle \phi_{\underline{\underline{\mathbf{R}}}}(t)| - i\partial_\alpha|\phi_{\underline{\underline{\mathbf{R}}}}(t)\rangle_{\underline{\mathbf{r}}}, \qquad (55)$$

and α is a composite index to denote the nucleus that is displaced and the cartesian direction of the displacement.

Explicitly, allowing to mix the different notations, the second term of the left-hand side of Eq. (53) becomes

$$\sum_{\mathbf{j}'} \; {}^{bn}\overline{H}_{\mathbf{j}}^{\mathbf{j}'}(t)p_{\mathbf{j}'}(t) = \sum_{\mathbf{ij}'} \pi_{\mathbf{ij}}^*(t){}^{bn}H_{\mathbf{j}}^{\mathbf{j}'}(t)\pi_{\mathbf{ij}'}(t)p_{\mathbf{j}'}(t)$$
$$= \langle \phi_{\underline{\underline{\mathbf{R}}}}(t)|\hat{T}_n(\underline{\underline{\mathbf{R}}})|\phi_{\underline{\underline{\mathbf{R}}}}(t)\rangle_{\underline{\mathbf{r}}} \; \chi(\underline{\underline{\mathbf{R}}},t)$$
$$= \sum_\nu \frac{1}{M_\nu}\left(-i\nabla_\nu + \mathbf{A}_\nu(\underline{\underline{\mathbf{R}}},t)\right)^2 \chi(\underline{\underline{\mathbf{R}}},t)$$
$$+ \Delta(\underline{\underline{\mathbf{R}}},t)\chi(\underline{\underline{\mathbf{R}}},t) \qquad (56)$$

where

$$\Delta(\underline{\underline{\mathbf{R}}},t) = \langle \phi_{\underline{\underline{\mathbf{R}}}}(t)| \sum_\nu \frac{1}{M_\nu}\left(-i\nabla_\nu - \mathbf{A}_\nu(\underline{\underline{\mathbf{R}}},t)\right)^2|\phi_{\underline{\underline{\mathbf{R}}}}(t)\rangle_{\underline{\mathbf{r}}}. $$
$$(57)$$

The scalar potential term, Eq.(9) of Ref. [12] is thus

$$\epsilon(\underline{\underline{\mathbf{R}}},t) = \epsilon_{\mathbf{j}}(t) + \Delta(\underline{\underline{\mathbf{R}}},t). \qquad (58)$$

With this definition, Eqs. (51) becomes Eq. (6) of Ref. [12], and Eq. (53) becomes Eq. (7) of Ref. [12].

Taking into account that the ${}^bH$ Hamiltonian is a laplacian, which is short-ranged in real space, there is a simple feedback from Eq. (18) to Eq. (42), through the gradient of $p_{\underline{\underline{\mathbf{R}}}}$, that is $p_{\mathbf{j}}$ in real space.

Note that the vector potential is defined through a simple derivative in Eq. (55). This means that it cannot be identified with a given piece of the hamiltonian. Its existence is intimately linked to the fact that application of the laplacian to a product of functions also yields a mixed term, where both functions are derived only once. Therefore, one may define such a contribution in every hamiltonian in real space that contains a kinetic energy

operator. Instead, for a more general Hamiltonian, in which the application of the $^b H$ operator to a product of functions does not yield separate products of terms, there is *a priori* no equivalent to Eqs.(54-55). Moreover, the usefulness of such a definition in a general basis is not evident.

2.9 Generality and specificities of the exact factorization strategy

Secs. 2.1 to 2.8 have explored, in a general sense, the outcomes of the exact factorization of a wavefunction in a marginal amplitude and a conditional amplitude (Sec.2.1), be it for the time-independent case or for the time-dependent case (Sec.2.7). It might perhaps be interesting for arbitrary Hamiltonians, in the case separate approximations can be made for the marginal component and for the conditional component, and one is interested in the marginal component density (Sec.2.5). The issue of the feedback loop between the marginal and conditional amplitudes has also been examined (Sec.2.4). If the Hamiltonian is semi-diagonal the governing equations simplify, see Sec.2.6, and even more if the operator governing the marginal amplitude dynamics has a specific form, like the laplacian (Sec.2.8). In the next section we will apply this understanding to a simple electron-boson coupling model, and illustrate selected variations on the theme of exact factorization.

3 Application to an electron-boson coupling model

3.1 One-site model

We consider first a one-site electron-boson coupling hamiltonian:

$$H = \omega_0 b^\dagger b + \tilde{g} c^\dagger c (b^\dagger + b). \quad (59)$$

Here c is a fermion annihilation operator and b a boson annihilation operator. For zero electron we have simply the energy of the bosons. For one electron, the electron-boson eigenstate $|q\rangle$ is a linear combination [49,50],

$$|q\rangle = \sum_m \alpha_q^m |1, m\rangle, \quad (60)$$

where $|1, m\rangle$ are states of one electron and m bosons. The coefficients α_q^m are known analytically: the eigenfunctions are displaced solutions of the harmonic oscillator, that can be expressed readily in the basis of the undisplaced harmonic oscillator eigenfunctions [49,50].

The corresponding eigenenergies are

$$e_q = -\frac{\tilde{g}^2}{\omega_0} + \omega_0 q, \quad (61)$$

where q is an integer, running from 0 to ∞, that counts the number of bosons in the displaced basis.

In the basis of undisplaced eigenfunctions the hamiltonian reads:

$$H_m^{m'} = \delta_{m,m'} \omega_0 m + \tilde{g}[\sqrt{m'+1}\delta_{m,m'+1} + \sqrt{m'}\delta_{m,m'-1}]. \quad (62)$$

In the exact factorization context, this one-site model is however a dummy case: there is no dependence of the coefficient α_q^m on an electron index, which corresponds to $\pi = 1$.

3.2 Two-site Holstein model

To have something interesting, we must enlarge the space of states for the electron(s). To this end we put one electron on the two-site Holstein model [35,36], with different site energies, whose Hamiltonian reads

$$H = \sum_{i=1}^2 \varepsilon_i c_i^\dagger c_i + \omega_0 \sum_{i=1}^2 b_i^\dagger b_i - t(c_1^\dagger c_2 + c_2^\dagger c_1)$$
$$+ g \sum_{i=1}^2 (b_i^\dagger + b_i) c_i^\dagger c_i, \quad (63)$$

where $c_i^{(\dagger)}$ and $b_i^{(\dagger)}$ are the annihilation (creation) operators of electrons of energy ε_i and non-dispersing (Einstein) bosons of energy ω_0, respectively, g is the *on-site* electron-boson coupling strength, and t represents the hopping kinetic energy of electrons between the two sites. Such Hamiltonian, a paradigmatic example of electron-boson coupling, has been studied for decades, as it naturally arises in different contexts from simplifications of the full-complexity Hamiltonians [37,39,41,42].

In the present context of exact factorization, the two-site Holstein model will prove to be very rich. It showcases different exact factorizations approaches. In the symmetric case, one even gets a perfect factorization, in which a one-site model is separated from a quantum Rabi model that is then treated with an exact factorization approach.

3.3 Two-site model, the naive approach

As basis functions for this problem we might chose states $|i; mn\rangle$, where the first argument indicates on which site the electron sits, whereas the second pair of numbers indicates the number of bosons on site 1 and 2, respectively.

The matrix elements of the hamiltonian in this basis are

$$H_{i;mn}^{i';m'n'} = {}^c H_i^{i'} \delta_{mm'} \delta_{nn'} + {}^b H_{mn}^{m'n'} \delta_{ii'} + {}^{cb} H_{i;mn}^{i';m'n'}, \quad (64)$$

with

$${}^c H_i^{i'} = \varepsilon_i \delta_{ii'} - t(\delta_{i',2}\delta_{i,1} + \delta_{i',1}\delta_{i,2}) \quad (65)$$
$${}^b H_{mn}^{m'n'} = \omega_0 \delta_{mm'} \delta_{nn'} (m+n) \quad (66)$$
$${}^{cb} H_{i;mn}^{i';m'n'} = g\delta_{i,i'}[\delta_{i,1}\delta_{nn'}(\delta_{m,m'+1}\sqrt{m} + \delta_{m,m'-1}\sqrt{m+1})$$
$$+ \delta_{i,2}\delta_{mm'}(\delta_{n,n'+1}\sqrt{n} + \delta_{n,n'-1}\sqrt{n+1})]. \quad (67)$$

With respect to the general formalism in the previous section, here i corresponds to the vector $\mathbf{i}$ of above, and the couple (m, n) corresponds to $\mathbf{j}$.

In the first "naive" attempt to use the exact factorization approach, we consider the electron system to be determined conditionally by the boson system, similarly to the usual exact factorization approach of the electron-nuclei system. Thus we express the coefficients C_{imn} as

$$C_{imn} \equiv \pi_{imn} p_{mn}, \quad (68)$$

with

$$\sum_{mn} |p_{mn}|^2 = 1 \quad \text{and} \quad \sum_i |\pi_{imn}|^2 = 1 \quad \text{for all } mn. \quad (69)$$

The time-independent equation for p is

$$\sum_{m'n'} \overline{H}_{mn}^{m'n'} p_{m'n'} = E p_{mn}, \qquad (70)$$

where

$$\overline{H}_{mn}^{m'n'} = {}^c\epsilon_{mn}\delta_{mm'}\delta_{nn'} + {}^b\overline{H}_{mn}^{m'n'} + {}^{cb}\overline{H}_{mn}^{m'n'}, \qquad (71)$$

with

$$\begin{aligned}
{}^c\epsilon_{mn} &= \varepsilon_1 \pi_{1mn}^* \pi_{1mn} + \varepsilon_2 \pi_{2mn}^* \pi_{2mn} \\
&\quad - t\big[\pi_{1mn}^*\pi_{2mn} + \pi_{2mn}^*\pi_{1mn}\big],
\end{aligned} \qquad (72)$$

$${}^b\overline{H}_{mn}^{m'n'} = \omega_0 \delta_{mm'}\delta_{nn'}(m+n), \qquad (73)$$

$$\begin{aligned}
{}^{cb}\overline{H}_{mn}^{m'n'} &= \\
&g\delta_{nn'}\ \big[\ \delta_{m,m'+1}\sqrt{m} + \delta_{m,m'-1}\sqrt{m+1}\big]\pi_{1mn}^*\pi_{1m'n'} \\
&+g\delta_{mm'}\ \big[\ \delta_{n,n'+1}\sqrt{n} + \delta_{n,n'-1}\sqrt{n+1}\big]\pi_{2mn}^*\pi_{2m'n'}.
\end{aligned} \qquad (74)$$

The effective potential corresponding to Eq. (23) is thus

$$\epsilon_{mn} = {}^c\epsilon_{mn} + \omega_0(m+n), \qquad (75)$$

without contribution from ${}^{cb}\overline{H}$. Supposing the π_{imn} values to be given, the reduction of the electron-boson problem to a boson problem with an effective potential, and effective interaction Eq. (74) is indeed achieved.

However, the equation for π is not simplified, as it includes off-diagonal matrix elements with respect to the (m,n) basis

$${}^{cb}h_{i;mn}^{i';m'n'} = \frac{p_{m'n'}}{p_{mn}}{}^{cb}H_{i;mn}^{i';m'n'}, \qquad (76)$$

where ${}^{cb}H_{i;mn}^{i';m'n'}$ is defined in Eq. (67). The time-dependent case is, of course, as complicated. We miss the semi-diagonal property. The two cases where it appears are presented in the next subsections.

3.4 Two-site model, exact factorization in position/momentum space

The Einstein oscillator sub-Hamiltonians can be expressed in the position/momentum space instead of using creation and annihilation operators. Defining

$$Q_i = \frac{1}{2}(b_i^\dagger + b_i) \quad \text{and} \quad P_i = \frac{i}{2}(b_i^\dagger - b_i) = -i\frac{\partial}{\partial Q_i}, \quad (77)$$

the two-site Holstein Hamiltonian becomes

$$\begin{aligned}
H &= \sum_{i=1}^2 \varepsilon_i c_i^\dagger c_i + \frac{\omega_0}{2}\sum_{i=1}^2 (P_i^2 + Q_i^2 - 1) - t(c_1^\dagger c_2 + c_2^\dagger c_1) \\
&\quad + g\sum_{i=1}^2 Q_i c_i^\dagger c_i.
\end{aligned} \qquad (78)$$

The total, conditional and marginal amplitudes write now

$$C_i(Q_1, Q_2) \equiv \pi_i(Q_1, Q_2)p(Q_1, Q_2). \qquad (79)$$

The Hamiltonian Eq.(78) is semi-diagonal with respect to the (Q_1, Q_2) coordinates that characterize the marginal amplitude $p(Q_1, Q_2)$. Apart from the choice of units, this situation is the same as for the usual real-space approach for electrons and nuclei, outlined in subsection 2.8, with the kinetic operator for the marginal amplitudes being $\frac{\omega_0}{2}\sum_{i=1}^2 P_i^2$. A vector potential

$$A_j(Q_1, Q_2) = \sum_{i=1}^2 \pi_i^*(Q_1, Q_2)(-i\frac{\partial}{\partial Q_j})\pi_i(Q_1, Q_2) \quad (80)$$

is introduced naturally. There is no penalty in having a discretized Hamiltonian for the conditional amplitude. The treatments and approximations introduced in the electron-nuclei case [20,21,51] can be followed. The treatment of the $e \otimes E$ Jahn-Teller model [25,26] is similar to this position/momentum treatment of the two-site Holstein model.

3.5 Two-site model, reverse exact factorization

If one is interested mainly in the electron density, the form of the cb coupling term Eq. (67) is already the desired semi-diagonal one, since it is diagonal in the i, i' indices. Now the boson component is associated with the conditional amplitude, as emphasized in Eqs. (42-44). In this case, the exact factorization starts from

$$C_{imn} \equiv \pi_{imn}p_i. \qquad (81)$$

with

$$\sum_i |p_i|^2 = 1 \qquad \text{and} \qquad \sum_{mn}|\pi_{imn}|^2 = 1 \quad \text{for all } i. \qquad (82)$$

In the case of the electron-nuclei system, Suzuki et al [15] proposed this alternative possibility that they called 'reverse exact factorization'.

The equations for the marginal amplitude (the electrons) are rather simple:

$$\epsilon_1 p_1 + {}^c\overline{H}_1^2 p_2 = E p_1 \qquad (83)$$

$$\epsilon_2 p_2 + {}^c\overline{H}_2^1 p_1 = E p_2 \qquad (84)$$

with

$${}^c\overline{H}_1^2 = {}^c\overline{H}_2^1 = -t\sum_{mn}\big[\pi_{1mn}^*\pi_{2mn} + \pi_{2mn}^*\pi_{1mn}\big]. \qquad (85)$$

Now, as expected, the equations governing π_{1mn} and π_{2mn} are also reasonably simple. The first one is

$$\begin{aligned}
\big(\varepsilon_1 &+ \omega_0(m+n)\big)\pi_{1mn} - t\frac{p_2}{p_1}(\pi_{2mn} - \pi_{1mn}) \\
&+ g\sum_{m'}[\delta_{m,m'+1}\sqrt{m} + \delta_{m,m'-1}\sqrt{m+1}]\pi_{1m'n} \\
&\qquad\qquad\qquad\qquad\qquad = \epsilon_1\ \pi_{1mn}\ . \quad (86)
\end{aligned}$$

The equation for π_{2mn} is similar, with 1 replaced by 2, and the sum running over n' instead of m', with replacement of m by n and m' by n' within the summation.

Without $\frac{p_2}{p_1}\pi_{2mn}$ this equation would have no coupling between 1 and 2, while the dependence on n would be trivial. Its eigenfunctions would be those of the one-site electron-boson coupling Hamiltonian, Eq. (62), which are analytically known.

Working now at fixed $\frac{p_2}{p_1}$, we might expand π_{1mn} and π_{2mn} in terms of such eigenfunctions α_q^m, which is an operation analogous to the Lang-Firsov transform [49],

$$\pi_{1mn} = \sum_q c_{1qn}\alpha_q^m$$
$$\pi_{2mn} = \sum_q c_{2qn}\alpha_q^m, \tag{87}$$

then pre-multiply Eq. (86) by $\alpha_{q'}^m$ and sum over m to get

$$c_{1qn} = \frac{t(p_2/p_1)c_{2qn}}{\epsilon_q + \omega_0 n + \varepsilon_1 - \epsilon_1 + t(p_2/p_1)}. \tag{88}$$

Here, ϵ_q is to be determined, under the constraint $\sum_{qn}|c_{1qn}|^2 = 1$. A similar equation and associated constraint exists for c_{2qn}.

In the reverse factorization case, the equations for p, π_1 and π_2, reveal clearly the special structure that has been discussed in Sec. 2.6, and as a consequence, the p appear in the equation for the conditional amplitude only in the kinetic term (proportional to t), as one can see from Eq. (86). The mathematical structure of the problem for each component of π is the one exemplified by Eq. (36), analyzed in Appendix B, also including the case of the coupling between such equations. Although complete decoupling can be obtained at zero t, when t is small, care should be taken to address equations for π_1 and π_2 in a joint approach, otherwise the procedure will not converge.

We will actually not address the solution of the joint conditional π_1-π_2 and marginal p equations for the reverse factorization treatment of the general two-site Holstein model, but will now consider specificities of the symmetric site case, for which we will provide a numerical analysis, including the interplay between the conditional and marginal amplitude equations.

3.6 Symmetric two-site Holstein model in the basis of natural orbitals, a perfect exact factorization

Since our approach is completely general, so is the definition of "the density": actually, we have defined it as the diagonal of the density matrix of the marginal amplitude, independently of the space. This implies that if we work in the basis of natural orbitals, where the density matrix is diagonal, our "density" at coordinate j corresponds to the occupation number of the natural orbital j.

We suppose $\epsilon_1 = \epsilon_2 = 0$ in the Holstein two-site model, where the electron wavefunction is considered to be the marginal amplitude. Due to symmetry reasons, the electronic natural orbitals are then the bonding (symmetric) and antibonding (antisymmetric) ones. We also choose a similar symmetrized or antisymmetrized basis for the bosonic degrees of freedom. The

corresponding creation and annihilation operators are

$$c_\pm^{(\dagger)} = \frac{1}{\sqrt{2}}\left(c_1^{(\dagger)} \pm c_2^{(\dagger)}\right); \tag{89a}$$
$$a_\pm^{(\dagger)} = \frac{1}{\sqrt{2}}\left(a_1^{(\dagger)} \pm a_2^{(\dagger)}\right), \tag{89b}$$

where the subscripts $+$ and $-$, respectively, represent *bonding* and *anti-bonding* orbitals, respectively.

The basis functions for the whole system are denoted by $|j, \ell, m\rangle$, where j indicates whether the electron is in the bonding ($j = 1$) or anti-bonding ($j = 2$) orbital, ℓ denotes the number of bonding bosons and m the number of anti-bonding bosons. The ground state wavefunction is developed on this basis, with coefficients $\mathcal{C}(j, \ell, m)$. [1]

In this basis, the Holstein Hamiltonian factorizes exactly as

$$H = H^{a+} + H^{a-} \tag{90}$$

with

$$H^{a+} = \omega_0 a_+^\dagger a_+ + \tilde{g}(c_+^\dagger c_+ + c_-^\dagger c_-)(a_+^\dagger + a_+); \tag{91a}$$
$$H^{a-} = -t(c_+^\dagger c_+ - c_-^\dagger c_-) + \omega_0 a_-^\dagger a_-$$
$$+ \tilde{g}(c_+ c_-^\dagger + c_- c_+^\dagger)(a_-^\dagger + a_-), \tag{91b}$$

where we have defined

$$\tilde{g} = g(\sqrt{2})^{-1} \tag{92}$$

.

In Eq. (91b) the $H^{a\pm}$ are separated according to the boson operators $a_\pm$. The two operators commute, and therefore

$$E = E^{a+} + E^{a-}; \tag{93a}$$
$$\mathcal{C}(j, \ell, m) = C^{a+}(\ell)C^{a-}(j, m): \tag{93b}$$

the ground state energy is the sum of the two ground state energies, and the wavefunction is the product wavefunction. In particular, a density $n(x)$ is simply the product of densities:

$$n(x) = \sum_{j\ell m}' |C^{a+}(\ell)|^2 |C^{a-}(j, m)|^2, \tag{94}$$

where x is a subset of the coordinates $(j\ell m)$ and $\sum_{j\ell m}'$ denotes the fact that the remaining coordinates are integrated out. For example, $n(j)$ is the electronic density in the bonding or anti-bonding orbital. Note that C^{a+} depends only on ℓ, whereas C^{a-} is independent of ℓ.

This is an exact factorization, and it is perfect, in the sense that the two problems can be solved completely independently.

The calculation of the ground state of H^{a+} is trivial, because the coupling term vanishes for one electron, and the lowest energy is obtained for zero boson occupation, with $C^{a+}(\ell) = \delta_{\ell,0}$. This means that the interesting densities are

$$n^{\text{el}}(j) \equiv \sum_m |C^{a-}(j, m)|^2, \tag{95}$$

---

[1] Note that what follows also holds for excited states, which would carry an extra label ν. One would simply have to order states according to the sum of two energies that are introduced below.
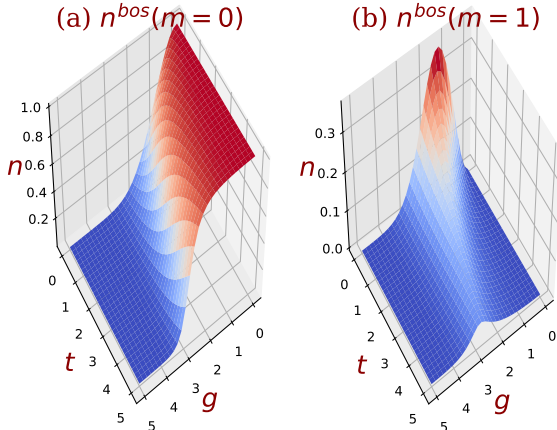
Figure 1. The densities $n^{bos}(m)$ for the case of $m = 0$ and $m = 1$. Here we set $\omega_0 = 1.0$.

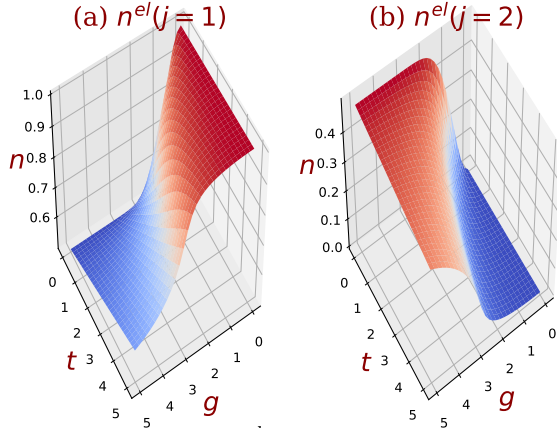**Figure 2.** The densities $n^{el}(j)$ for the case of bonding $j = 1$ and antibonding orbital $j = 2$. Here we set $\omega_0 = 1.0$.

and

$$n^{\text{bos}}(m) \equiv \sum_j |C^{a-}(j,m)|^2. \qquad (96)$$

This requires the calculation of the ground state of H^{a-}, which is the interesting part here. This Hamiltonian is also the one of the quantum Rabi model, that, like the initial (possibly unsymmetric) Holstein model, has a long history, and is also often referred to as the Hamiltonian of the two-level pseudo Jahn-Teller effect, see e.g. Refs. [37,38,40,41,43].

3.7 The quantum Rabi Hamiltonian

We now focus on an exact factorization treatment of H^{a-}, the quantum Rabi Hamiltonian. For simplicity, we will omit the superscript, $C^{a-}(j,m) \to C(j,m)$, and similarly, $H^{a-} \to H$.

H is first expressed in the basis $|j,m\rangle$, where the notation is defined as for $|j,\ell,m\rangle$ above:

$$H_{jm}^{j'm'} = \delta_{jj'}\delta_{mm'}(-t(\delta_{j,1} - \delta_{j,2}) + m\omega_0)$$
$$+\tilde{g}\ (\delta_{j',1}\delta_{j,2} + \delta_{j,1}\delta_{j',2})(\delta_{m,m'+1}\sqrt{m} + \delta_{m,m'-1}\sqrt{m'}). \qquad (97)$$

At this point, we have two choices. The first one is a straightforward application of the exact factorization procedure. Thus, we have

$$C(j,m) = p_m\pi_{jm} \qquad (98)$$

for the exact factorization, where the aim is to calculate the density of bosons, and

$$C(j,m) = p_j\pi_{jm} \qquad (99)$$

for the exact reverse factorization, where the aim is to calculate the density of electrons.

Let us start with the boson density, $n^{\text{bos}}(m) = \sum_j |C(j,m)|^2$. The first two components ($m = 0$ and $m = 1$) are shown in Fig. 1 for a fixed value of $\omega_0 = 1.0$, as a function of the coupling constant g and the hopping t. For vanishing coupling all the density is found in the $m = 0$ component. With increasing g weight is shifted to the $m = 1$ component, which has a maximum between $g = 1$ and $g = 3$ for the values of t shown here. For even stronger coupling both the $m = 1$ and $m = 0$ components decrease, because higher components gain in weight. With increasing t, this behavior is shifted to higher values of g, since it becomes more difficult to couple the levels.

For the factorization, we define

$$p_m \equiv \sqrt{\sum_j |C(j,m)|^2}e^{\alpha(m)} \qquad (100)$$

and

$$\pi_{jm} \equiv C(j,m)/p_m \qquad (101)$$

and obtain the effective Schrödinger equation for p_m,

$$\bar{H}_m^{m'}p_{m'} = Ep_m \qquad (102)$$

with

$$\bar{H}_m^{m'} = \sum_{jj'} \pi_{jm}^* H_{jm}^{j'm'} \pi_{j'm'}$$
$$= \delta_{mm'}\left(t(|\pi_{2m}|^2 - |\pi_{1m}|^2) + m\omega_0(|\pi_{2m}|^2 + |\pi_{1m}|^2)\right)$$
$$+ \tilde{g}(\pi_{2m}^*\pi_{1m'} + \pi_{1m}^*\pi_{2m'}) \times$$
$$\times (\delta_{m,m'+1}\sqrt{m} + \delta_{m,m'-1}\sqrt{m'}). \qquad (103)$$

The effective hamiltonian for the π_{jm} relies on the computation of $\frac{1}{p_m}H_{jm}^{j'm'}p_{m'}$. It would be tempting to start from the non-interacting result obtained with $\tilde{g} = 0$, which is $p_m = \delta_{m,0}$, as can be seen from Fig. 1, or $\pi_{jm} = \delta_{j,1}$. However, because of the division by p_m this is immediately problematic.

If we are heading for $n_j^{\text{el}} = \sum_m |C(j,m)|^2$, we will opt for the reverse factorization. The exact result is shown in Fig. 2.There are only two possible components corresponding to the bonding ($j = 1$) and antibonding ($j = 2$) orbitals. As expected, the bonding orbital is fully occupied for $g = 0$, and with increasing g, weight is shifted to the antibonding orbital. As in the case of the boson occupations, this interaction effect occurs later when t is larger.

For the factorization, we define

$$p_j \equiv \sqrt{\sum_m |C(j,m)|^2}e^{\alpha(j)} \qquad (104)$$

and

$$\pi_{mj} \equiv C(j,m)/p_j \tag{105}$$

and obtain the effective Schrödinger equation for p_j,

$$\bar{H}_j^{j'} p_{j'} = E p_j \tag{106}$$

with

$$\begin{aligned}
\bar{H}_j^{j'} &= \sum_{mm'} \pi_{mj}^* H_{mj}^{m'j'} \pi_{m'j'} \\
&= \delta_{jj'} \left(t\, s(j) + \sum_m m |\pi_{mj}|^2 \omega_0 \right) \\
&\quad + \tilde{g}(\delta_{j',1}\delta_{j,2} + \delta_{j,1}\delta_{j',2}) \times \\
&\quad \times \sum_m (\sqrt{m}\, \pi_{mj}^* \pi_{m-1,j'} + \sqrt{m+1}\, \pi_{mj}^* \pi_{m+1,j'}),
\end{aligned} \tag{107}$$

where $s(j) = -1$ for $j = 1$ and $s(j) = 1$ for $j = 2$.

The effective hamiltonian for the π_{mj} also relies on $\frac{1}{p_j} H_{jm}^{j'm'} p_{j'}$. Again, it would be tempting to start from the non-interacting result obtained with $\tilde{g} = 0$, which is here $p_j = \delta_{j,1}$ or $\pi_{mj} = \delta_{m,0}$ (note the perfect inversion of the role of the two parameters with respect to the exact factorization above). However, now because of the division by p_j this is again problematic.

3.8 Iterative exact factorization treatment of the ground state of the quantum Rabi model

The hamiltonian (97) shows that states with odd j (bonding) and even m couple only to states with even j (antibonding) and odd m, and similarly for states with even-j (anti-bonding) and even m states that couple only to states with odd-j (anti-bonding) and odd m states [37]. Thus, one can focus on one of these subspaces only (which constitutes one more exact perfect factorization), actually the 'bonding+even-m plus antibonding+odd-m' subspace, as the ground state of the system belongs to it [40]. Thanks to this insight, the initialization problem that we faced in the previous subsection disappears. For sake of clarity, we will not modify our notations to indicate that, for fixed j, the summation over m runs only on one subset of these indices, as this can stay implicit without confusion.

Eq. (106) is a two by two matrix equation for the marginal amplitudes. Its conditional amplitude dependence comes from

$$\begin{aligned}
{}^b\overline{H}_1^1 &= \sum_m m |\pi_{m1}|^2 \omega_0, \\
{}^b\overline{H}_2^2 &= \sum_m m |\pi_{m2}|^2 \omega_0, \\
{}^{cb}\overline{H}_1^2 &= \tilde{g} \sum_m (\sqrt{m}\, \pi_{m1}^* \pi_{m-1,2} + \sqrt{m+1}\, \pi_{m1}^* \pi_{m+1,2}), \\
{}^{cb}\overline{H}_2^1 &= \tilde{g} \sum_m (\sqrt{m}\, \pi_{m2}^* \pi_{m-1,1} + \sqrt{m+1}\, \pi_{m2}^* \pi_{m+1,1}).
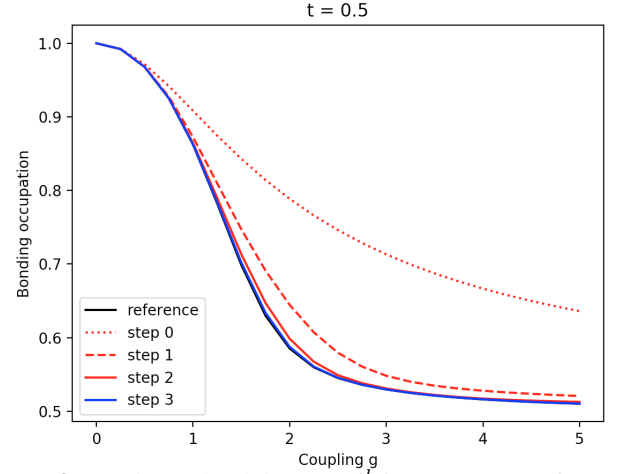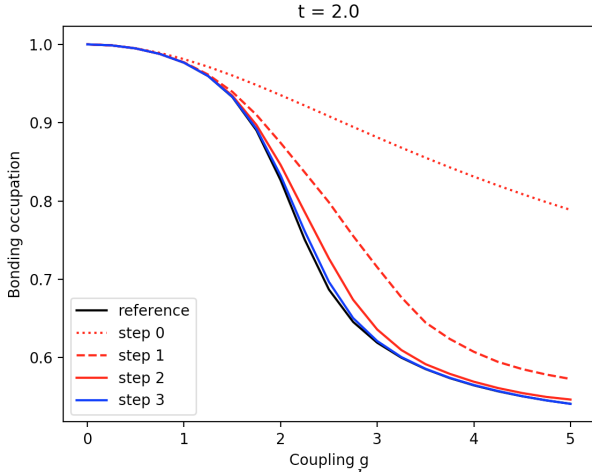\end{aligned} \tag{108}$$

Figure 3. Bonding orbital density n^{el} for $t = 0.5$, as a function of g (with $\omega_0 = 1$) at different steps of the exact factorization iterative procedure. Reference exact result, continuous black line; initialization of the exact factorization (step 0), dotted red line; first, second and third step of the exact factorization, dashed red, continuous red and continuous blue lines. Step 3 and reference can hardly be distinguished.

Note that ${}^b\overline{H}$ is j-diagonal, while ${}^{cb}\overline{H}$ has no j-diagonal component. Like ${}^b\overline{H}$, ${}^c\overline{H}$ is j-diagonal,

$$ {}^c\overline{H}_j^{j'} = \delta_{jj'} t\, s(j). \tag{109}$$

Following the discussion in Sec. 2.4, we will illustrate the determination of the marginal probability, using an iterative procedure. Working now in the proper subspace, we start the iterative process with $\pi_{01} = 1$ and $\pi_{12} = 1$, with all other conditional amplitudes being zero. This trivially fulfills the partial normalization conditions. Solving the two by two marginal equation is trivial, which delivers initial (step=0) values for the marginal amplitudes. The corresponding bonding and anti-bonding densities are computed from Eq. (95), and represented in Figs. 3-5, as a function of the g parameter (see its relationship with $\tilde{g}$, Eq.92) , for three fixed values of t, namely $t = 0.5$, 2.0 and 5.0, in the case $\omega_0 = 1$. The exact result, from Fig. 2, is also represented (i.e. cuts in the left part of Fig. 2 with $t = 0.5$, $t = 2.0$ and $t = 5.0$).

This simple initialization delivers bonding orbital occupations in excellent agreement with the reference results in the range where the contribution from boson numbers two and higher is small, namely up to $g \approx 0.6$ for $t = 0.5$, up to $g \approx 1$ for $t = 2.0$, and up to $g \approx 2$ for $t = 5.0$. For larger values of g, the bonding occupation departs significantly from the reference.

We then proceed with the iterative process. During the first step, the occupation of bosons is optimized at fixed values of p_i, using the preconditioned conjugate gradient algorithm [47,48], with precondition gradients defined by Eq. 40, and a diagonal preconditioner. An arbitrary number of bosons is then included, be it in the even-boson channel or in the odd-boson channel (in practice, we fixed the maximum number of bosons to 25, which guarantees 6 significant digits on the energy). The bonding occupation number is then recomputed, using updated $\overline{H}$ matrix elements. This delivers the results dashed-red 'step 1' results

Figure 4. Bonding orbital density n^{el} for $t = 2.0$, as a function of g (with $\omega_0 = 1$) at different steps of the exact factorization iterative procedure. Labeling is the same as in Fig. 3. Step 3 and reference differ slightly for g in the 2.0-3.0 range.
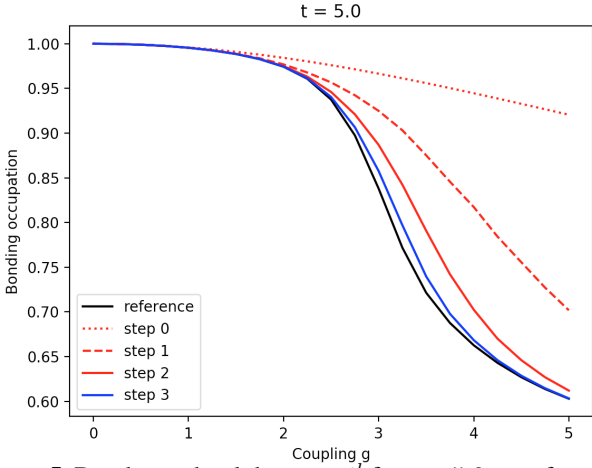

Figure 5. Bonding orbital density n^{el} for $t = 5.0$, as a function of g (with $\omega_0 = 1$) at different steps of the exact factorization iterative procedure. Labeling is the same as in Fig. 3. Step 3 and reference differ somehow for g in the 3.0-4.0 range.

in Figs. 3-5, with a significantly improved prediction. The subsequent second iteration delivers results within a few percent agreement, while the third iteration is nearly indistinguishable from the reference in the $t = 0.5$ and $t = 2.0$, but still differs visually in the $t = 5.0$ case.

Of course, this quantum Rabi model is particularly simple, and any usual matrix treatment delivers easily the reference results. The emphasis was laid here simply on the illustration of a flow in which the two subsystems are treated separately using exact factorization.

3.9 Quantum Rabi model, exact factorization coupled with perturbation theory

Since without further approximations the equation for π is as hard to solve as the full Schrödinger equation, it is clear that an approximation has to be done. As an example, we will in the

following discuss a perturbative approach, where the coupling is treated perturbatively in the equation for the π, but not for the p.

In principle, to do so one would start from the equation for π_{mj} and solve it to a given order in the interaction. However, this equation is not an ordinary Schrödinger equation; in particular, the normalization condition applies for each j separately. Therefore it is dangerous to straightforwardly apply the usual expressions of perturbation theory. Instead, we start from the original hamiltonian Eq. (97), for which we can do ordinary perturbation theory in $\tilde{g}$. The zero order eigenvalues and the coefficients of the eigenfunctions are (ν labels different states and $\nu = 0$ corresponds to the ground state):

$$E_\nu^{(0)} = ts(j_\nu) + m_\nu\omega_0 \,, \qquad (110)$$

$$C_\nu^{(0)}(j, m) = \delta_{j,j_\nu}\delta_{m,m_\nu} \,. \qquad (111)$$

For the ground state, the results are $E^{(0)} = -t$ and $C^{(0)}(j, m) = \delta_{j,1}\delta_{m,0}$. The corresponding p's are calculated from their definition, leading to $p_{j=1}^{(0)} = 1$ and $p_{j=2}^{(0)} = 0$. Using these zero order elements and again the definition of p, we obtain for the first order ground state coefficients

$$C^{(1)}(j, m) = -\frac{\tilde{g}}{\omega_0+2t}\delta_{j,2}\delta_{m,1} \,, \qquad (112)$$

$$p_j^{(1)} = \frac{\tilde{g}}{\omega_0+2t}\delta_{j,2} \,. \qquad (113)$$

From this, one would expect to obtain a first-order expression for the π_{mj}, but

$$\pi_{mj} \approx \frac{C^{(0)}(j, m) + C^{(1)}(j, m)}{p^{(0)} + p^{(1)}} \qquad (114)$$

leads to

$$\pi_{mj} = \delta_{j,1}\delta_{m,0} + \delta_{j,2}\delta_{m,1}\frac{-\tilde{g}/(2t + \omega_0)}{\tilde{g}/(2t + \omega_0)}$$
$$= \delta_{j,1}\delta_{m,0} - \delta_{j,2}\delta_{m,1} : \qquad (115)$$

the first-order dependence on $\tilde{g}$ cancels, and a zero-order result is obtained. This illustrates again the fact that one has to be careful when dealing with perturbation theory here.

Our starting π_{mj} hence defined from perturbation theory is, besides an (arbitrary) phase, the same as the starting point of the exact iteration in the previous subsection. Therefore it leads to the same effective hamiltonian for the p_j, which is

$$\bar{H}_j^{j'} = \delta_{jj'}(ts(j) + \delta_{j,2}\omega_0) - \tilde{g}(\delta_{j',1}\delta_{j,2} + \delta_{j',2}\delta_{j,1}). \quad (116)$$

It yields the ground state energy

$$E = \frac{\omega_0}{2} - \sqrt{(\frac{\omega_0}{2} + t)^2 + \tilde{g}^2}, \qquad (117)$$

which is consistent with the ordinary perturbation result $E \approx -t - \tilde{g}^2/(2t + \omega_0)$. The coefficients are

$$p_1^2 = \frac{1}{2}(1 + \frac{1}{\sqrt{1 + x^2}}) \approx 1 - \frac{\tilde{g}^2}{(2t + \omega_0)^2}$$

$$p_2^2 = 1 - p_1^2 \approx \frac{\tilde{g}^2}{(2t + \omega_0)^2},$$

$$\text{with} \qquad x^2 \equiv \frac{4\tilde{g}^2}{(\omega_0 + 2t)^2} : \qquad (118)$$
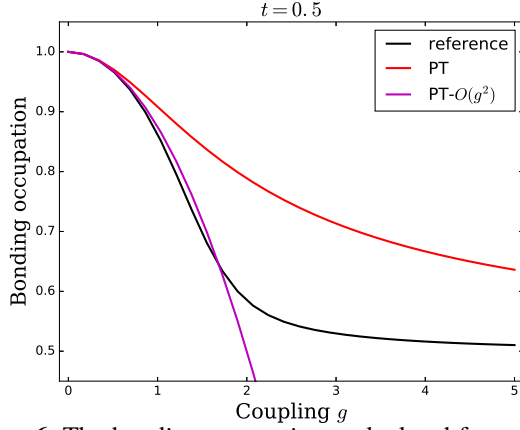
Figure 6. The bonding occupations calculated from perturbation theory. The red curve is p_1^2 in equation (118) and the magenta curve is p_1^2 in order $O(g^2)$. The black and red curves are the same as the black and dotted red line in figure 3. Interestingly, the $O(g^2)$ result outperforms the one of (118) for weak coupling.

Figure 7. The bonding occupations calculated from the perturbation theory. The red curve is p_1^2 in equation (118) and the magenta curve is p_1^2 in order $O(g^2)$. The black and red curves are the same as the black and dotted red line in figure 4.

Figure 8. The bonding occupations calculated from the perturbation theory. The red curve is p_1^2 in equation (118) and the magenta curve is p_1^2 in order $O(g^2)$. The black and red curves are the same as the black and dotted red line in figure 5.

the full expressions (first expression on the right hand side) are the same as the outcome of the "step 0" in the previous subsection. They are shown again in Fig.s 6 to 8, together with their expansion to first order in $\tilde{g}$. The latter quickly diverges from the exact result for strong coupling, while it yields decent results for weak to moderate coupling, even better than the full expressions. The comparison of various perturbative results is interesting here since, contrary to Subsec. 3.7, the expressions are not merely used as input for further iteration of the exact problem: the aim is rather to discuss them from the point of view of a possible approximation scheme, defined as: "Determine π to a given order in perturbation theory" (here, order zero), and "solve the resulting equation for p exactly or approximately". Even in case of an exact solution of the approximate $\bar{H}$, this reduces the work with respect to the exact solution of the initial problem, contrary to the full iterative approach in the previous subsection. The interesting question to be discussed here is how the final result compares with what one could obtain with a similar effort by approximating directly the initial problem.

Here, the workload has been to first obtain the π in first order perturbation theory on the initial problem. The direct outcome of this perturbative step, as one can see from (113), equals the $O(\tilde{g}^2)$-expansion of the p_2 that is obtained in the factorization method. The latter adds higher orders through the diagonalization of $\bar{H}$. For p_1, there is no first-order correction from the initial perturbation calculation; a correction would only appear in second order perturbation theory, such restoring the correct normalization of the p^2. The factorization result is therefore much superior. This might be linked to the fact that the p are directly square roots of the densities, instead of wavefunctions, which combine different orders to yield the density at a given order. The price to pay is the diagonalization of $\bar{H}$; when, as it is the case here, this hamiltonian is much smaller than the initial one, the factorization approach is an efficient choice.
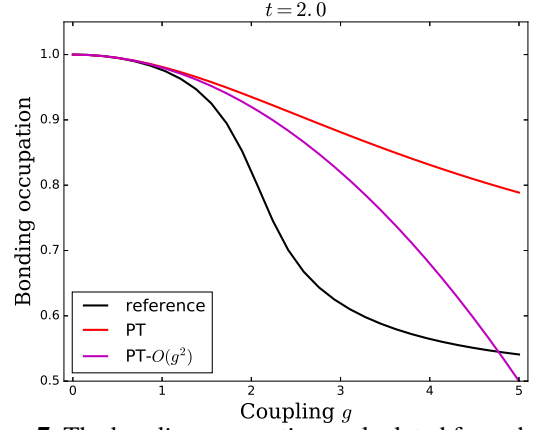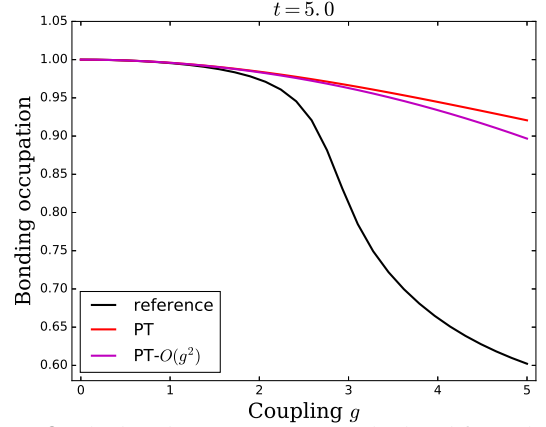
3.10 Factorization or frequency-dependent effective hamiltonian?

As explained at the beginning of Subsec. 3.7, the knowledge of m for the ground state determines whether it corresponds to the bonding or anti-bonding state. In the following, we work with $C(j, m) \rightarrow c(m)$, understanding that $c(m)$ with $m = $ even is $C(1, m)$ and $c(m)$ with $m = $ odd is $C(2, m)$. The effective hamiltonian $H_m^{m'}$ resulting from Eq. 97 is

$$H_m^{m'} = \delta_{mm'}(m\omega_0 + t\,s(m))$$
$$+ \tilde{g}(\delta_{m'+1,m}\sqrt{m} + \delta_{m',m+1}\sqrt{m'}), \qquad (119)$$

where

$$s(m) = 1 \quad \text{for } m \text{ odd}, \quad \text{and} \quad -1 \quad \text{for } m \text{ even}. \quad (120)$$

How are now the densities calculated? The boson density is simply $n^{\text{bos}}(m) = |c(m)|^2$: all knowledge about the electron-boson coupling has already been used, and nothing more can be

done. The electron density, instead, is $n^{\text{el}}(j=1) = \sum_{m\,\text{even}} |c(m)|^2$ and $n^{\text{el}}(j=2) = \sum_{m\,\text{odd}} |c(m)|^2$. This creates a very interesting situation: now, instead of summing over *one of the coordinates*, we sum over *a part of the range of the only one coordinate that is left*. In other words, out of the vector $c(m)$ we only want to know certain elements, according to whether we are interested in the bonding or the anti-bonding electron density.

To work this out, let us divide the hamiltonian $H_m^{m'}$ into four blocks, with indices 1 for m, m' even and 2 for m, m' odd. These blocks read:

$$H_m^{m'} =$$
$$= \delta_{mm'}(m\omega_0 - t) \qquad m, m' \text{ even}$$
$$= \delta_{mm'}(m\omega_0 + t) \qquad m, m' \text{ odd}$$
$$= \tilde{g}(\delta_{m'+1,m}\sqrt{m} + \delta_{m',m+1}\sqrt{m'}) \quad m \text{ even, } m' \text{ odd}$$
$$\text{or } m \text{ odd, } m' \text{ even.} \qquad (121)$$

Now one can use the following idea: Suppose we have a $(2n \times 2n)$ eigenvalue problem

$$\begin{pmatrix} \tilde{H}_1^1 & \tilde{H}_1^2 \\ \tilde{H}_2^1 & \tilde{H}_2^2 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \omega \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \qquad (122)$$

where $\tilde{H}_k^\ell$ are $n \times n$ matrices. Solving for $\phi_2 = (\tilde{H}_2^2 - \omega)^{-1} \tilde{H}_2^1 \phi_1$ and substituting into the equation for ϕ_1 one gets

$$[\tilde{H}_1^1 - \tilde{H}_1^2(\tilde{H}_2^2 - \omega)^{-1}\tilde{H}_2^1]\phi_1 = \omega\phi_1, \qquad (123)$$

provided that $(\tilde{H}_2^2 - \omega)$ is invertible. The latter is an $(n \times n)$ eigenvalue problem for a frequency-dependent matrix $\tilde{H}(\omega) = \tilde{H}_1^1 - \tilde{H}_1^2(\tilde{H}_2^2 - \omega)^{-1}\tilde{H}_2^1$. The eigenvalues obtained from Eq. (123) are the same as those of Eq. (122). Note that the number of eigenvalues corresponds to the one of the original matrix and is hence twice the dimension of the $(n \times n)$ matrix, which is possible due to the frequency dependence. Of course, part of the information concerning the eigenvectors of the original problem are lost, but this is not necessarily a problem when one is only interested in partial information.

In general the frequency dependent contribution in the new matrix, which plays the role of a self-energy, is not known. It contains all matrix elements of the original hamiltonian besides $\tilde{H}_1^1$, but since they are partially summed over, one may hope to find suitable approximations.

With the separation Eq. (121) we can directly apply this idea to our problem: this leads to two frequency-dependent effective hamiltonians, one to calculate even components, and another one to calculate odd components, and therefore the corresponding electron densities. For illustration, we concentrate on the even component. The self-energy reads

$$\Sigma_{mm'}(\omega) \equiv [-\tilde{H}_1^2(\tilde{H}_2^2 - \omega)^{-1}\tilde{H}_2^1]_{mm'}$$
$$= -\tilde{g}^2 \frac{\delta_{nn'}}{(n\omega_0 + t - \omega)} \times$$
$$\times (\delta_{n+1,m}\sqrt{m} + \delta_{n,m+1}\sqrt{n}) \times$$
$$\times (\delta_{m'+1,n'}\sqrt{n'} + \delta_{m',n'+1}\sqrt{m'}), \qquad (124)$$

where the repeated indices n, n' are summed over. This yields

$$\Sigma_{mm'}(\omega) = -\tilde{g}^2 \Big[\frac{m\,\delta_{m',m}}{(m-1)\,\omega_0 + t - \omega}$$
$$+ \frac{(m+1)\,\delta_{mm'}}{(m+1)\,\omega_0 + t - \omega} +$$
$$+ \frac{\delta_{m'+1,m-1}\sqrt{m}\sqrt{m'+1}}{(m'+1)\omega_0 + t - \omega} +$$
$$+ \frac{\delta_{m'-1,m+1}\sqrt{m+1}\sqrt{m'}}{(m+1)\omega_0 + t - \omega} \Big]. \qquad (125)$$

This expression is of first order in $\tilde{g}^2$. The self-energy for the odd component would be obtained by replacing everywhere $t \to -t$.

To obtain the desired density, we calculate the Green's function $G_{mm'}^{\text{even}}(\omega) = \left(G_0^{-1}(\omega) - \Sigma^{\text{even}}(\omega)\right)_{mm'}^{-1}$, where $G_{0,mm'}(\omega) \equiv \frac{\delta_{mm'}}{\omega - m\omega_0 + t - i\eta}$, and calculate the bonding density from the weights of the poles in the result. For simplicity and illustration, we use only the $m = m' = 0$ case of the Dyson equation, which is of course an approximation. Then we have

$$G_0(\omega) \qquad = \frac{1}{\omega + t - i\eta} \, ; \qquad (126)$$
$$\Sigma(\omega) \qquad = \frac{\tilde{g}^2}{\omega - t - \omega_0 - i\eta} \, . \qquad (127)$$
$$G(\omega) = (G_0(\omega)^{-1} - \Sigma(\omega))^{-1}, \qquad (128)$$
$$= \frac{z_1}{\omega - \varepsilon_1 - i\eta} + \frac{z_2}{\omega - \varepsilon_2 - i\eta} \, . \qquad (129)$$

The self-energy clearly exhibits the physics of the coupling: it has a pole at the bare anti-bonding energy augmented by one boson. Indeed, it can be written as a convolution of the anti-bonding component of the non-interacting G_0 and the boson propagator, and it is therefore identical to the anti-bonding contribution to the self-energy in the GW approximation in the case where the boson stems from electronic excitations, or the Dyson-Migdal approximation in the cased of phonons.

From this self-energy we obtain for the poles and spectral weight of the Greens' function G:

$$\varepsilon_1 = \frac{\omega_0}{2} - \sqrt{(\frac{\omega_0}{2} + t)^2 + \tilde{g}^2} \, ;$$
$$\varepsilon_2 = \frac{\omega_0}{2} + \sqrt{(\frac{\omega_0}{2} + t)^2 + \tilde{g}^2} \, ;$$
$$z_1 = \frac{1}{2}\Big(1 + \frac{1}{\sqrt{1 + \frac{4\tilde{g}^2}{(\omega_0 + 2t)^2}}}\Big) \, ;$$
$$z_2 = 1 - z_1 \, . \qquad (130)$$

Whereas G_0 has only one pole, the resulting spectral function, shown in Fig. 9, has two structures, one corresponding to the ground state energy, and one corresponding to an excited state that stems from the block taken into account by the self-energy. For small coupling, the ground state pole has the correct first order expansion $\omega = -t - \frac{\tilde{g}^2}{\omega_0 + 2t}$. Its weight, which gives the $m = 0$ component of the ground state boson density, is given by z_1 in (130). For vanishing coupling the bonding orbital is fully occupied and there is no boson. The coupling leads to excitation of bosons (even m and to a partial occupation of the anti-bonding orbital (odd m), and the effect is enhanced when the hopping t and the boson frequency ω_0 are small.
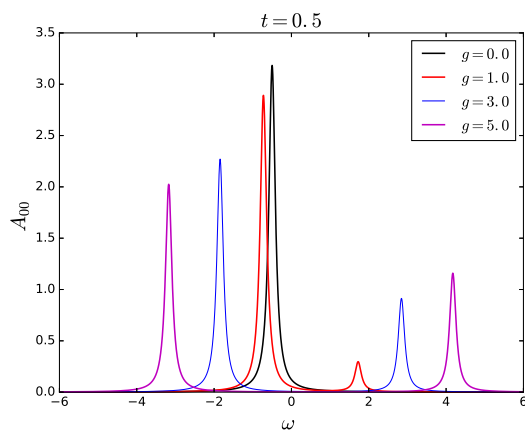
Figure 9. The spectral function from G_{00}. The intensity of the structure at lower energy indicates the occupation of the bonding orbital.

By examining the expressions one finds that interestingly, the weights z_1 and z_2 are the same as the first results for p_1^2 and p_2^2 in Eq. (118), i.e. the weights obtained from the factorization approach combined with initial perturbation theory. Also the workload is comparable: setting up the self-energy is equivalent to the initial perturbative step, followed by the diagonalization of a 2×2 matrix. The situation might however be specific to the present problem, and a more general analysis is needed to draw guidelines for the choice of methods.

Concerning the principles of the factorization versus the self-energy approach, it should be noted that the latter does not require the transformation of the problem to make the coordinate j disappear, as we have done here, but one could have applied it also to calculate the even boson number density at an earlier stage, when keeping still j. The important distinction between the factorization and the frequency-dependent formulation is the following:

The factorization is pertinent when one wants to integrate out one or more coordinates in the solution (in the wavefunction squared, to be precise). The frequency-dependent approach is pertinent when one asks for the wavefunction calculated at certain values of one or more coordinates only, whereas other values are not of interest.

4 Conclusion and perspectives

In conclusion, in this work we have tried to work out some fundamental features of the exact factorization approach on one side, and to elucidate and illustrate its functioning in practice, with the help of a model, on the other side. In order to do so we have essentially discussed the time-independent case, but most of our discussion is straightforwardly carried over to the time-dependent problem.

The exact factorization is an in principle exact rewriting of the many-body problem, for arbitrary hamiltonians. The electron-nuclei system, on which the work of Hardy Gross and coworkers puts its accent, belongs to a particular class of hamiltonians for which the effect of one species can be expressed as an additional local potential appearing in an effective Schrödinger equation

for an effective wavefunction of the other species. This effective wavefunction can be used to calculate the N-body density, but not necessarily other observables. Other hamiltonians will also allow one to use the effective wavefunction to calculate the N-body density of one species, but the other species may give a contribution to the effective hamiltonian that differs from a local potential. In any case, various possible particular types of hamiltonians suggest different ways to transform and approximate the equations. In the present work we have discussed these general points and briefly made the link to the specific electron-nuclei problem. In order to illustrate several points, we have used electron-boson coupling models, where results are compact and relatively easy to analyze.

From a pragmatic point of view, one may argue that any rewriting of the many-body problem is useful only if it suggests powerful approximations. Does the exact factorization approach introduce new openings in this sense? Since approximations are always to some extent system-specific, it is by definition not possible to give a completely general answer. However, the model study in our work shows some interesting features. On one side, both by performing the iteration of the coupled equations exactly and by partially applying perturbation theory, it appears that care is needed when dealing with the equations, for example, because of the particular normalization of the conditional amplitude. On the other side, it has been interesting to compare the approximate results with straightforward perturbation theory on one hand, and with a self-energy approach on the other hand. At least in the model studied here, the exact factorization scheme seems to be more efficient than straightforward perturbation theory, performed with a similar effort. We could not draw such a conclusion concerning the alternative of using a self-energy. However, we also explained that the factorization and the self-energy approach do not pose the question of interest in the same way: the factorization approach targets observables where a part of the variables of the wavefunction are integrated out, whereas the self-energy approach is intuitive when one asks for the value of a function in a part of the domain of its variables. This explains why the self-energy introduces a frequency-dependence, while the factorization doesn't. The particularities of the model studied here made this comparison more straightforward than what it might be in general.

Altogether, by trying to separate the electron-nuclei-specific features from fundamental characteristics of a factorization approach, we found that several interesting routes could be taken to explore the ideas in the framework of other particular problems. A general powerful approximation strategy for the factorization approach is probably still missing, but there is some evidence that the approach has the potential for such progress.

Acknowledgements

Author contribution statement

A Phase covariance and contravariance

In this short appendix, we mention the consequences of the phase invariance of the decomposition Eq. (2). Indeed, the solution of the Schrödinger equation Eq. (15) is left invariant upon the replacements

$$\pi_{\mathbf{ij}} \rightarrow \tilde{\pi}_{\mathbf{ij}} = e^{i\theta_{\mathbf{j}}}\pi_{\mathbf{ij}} \quad \text{and} \quad p_{\mathbf{j}} \rightarrow \tilde{p}_{\mathbf{j}} = e^{-i\theta_{\mathbf{j}}}p_{\mathbf{j}}, \qquad (131)$$

which corresponds to a different choice of phases $\alpha(\mathbf{j})$. The behavior of $\pi_{\mathbf{j}}$ will be referred to as phase-covariance, and the one of $p_{\mathbf{j}}$ will be referred to as phase-contravariance. The b and bc terms transform as

$$^{b}\overline{H}_{\mathbf{j}}^{\mathbf{j}'} \rightarrow {}^{b}\widetilde{\overline{H}}_{\mathbf{j}}^{\mathbf{j}'} = e^{i(\theta_{\mathbf{j}} - \theta_{\mathbf{j}'})}{}^{b}\overline{H}_{\mathbf{j}}^{\mathbf{j}'}, \qquad (132)$$

$$^{cb}\overline{H}_{\mathbf{j}}^{\mathbf{j}'} \rightarrow {}^{cb}\widetilde{\overline{H}}_{\mathbf{j}}^{\mathbf{j}'} = e^{i(\theta_{\mathbf{j}} - \theta_{\mathbf{j}'})}{}^{cb}\overline{H}_{\mathbf{j}}^{\mathbf{j}'}. \qquad (133)$$

As the final computed quantities are invariant with the phase choice, one is guided naturally to treat differently quantities that are phase invariant (like the diagonal of the b and bc terms) and those that have phase-dependence, which yields the definition Eq.(23). In Sec.2.8, the specific form of the ^{b}H Hamiltonian, a Laplacian, allows one to handle the phase transformations in a very elegant way, as emphasized in Refs.[10,12].

B Mathematical structure of the conditional amplitude effective equation

B.1 Preliminary analysis

Mathematically, the equation for the conditional amplitude $\pi_{\mathbf{ij}}$ in the time-independent case, see Eq. (35) or (42) does not correspond to a usual eigenvalue problem. The presence of the $\pi_{\mathbf{i}'\mathbf{j}'}$ or $\pi_{\mathbf{ij}'}$ term, with the normalization condition $\sum_{\mathbf{i}}|\pi_{\mathbf{ij}}|^2 = 1$, that made us invoke the Lagrange multiplier approach, turns it to an equation of the type

$$\underline{\underline{A}}\underline{x} + \underline{y} = \epsilon\underline{x} \qquad (134)$$

under constraint $\underline{x}^{\dagger}\underline{x} = 1$. Unlike in a Schrödinger equation, an inhomogeneous term $\underline{y}$ is present.

Such equation can be solved by expanding $\underline{x}$ and $\underline{y}$ in terms of the eigenvectors of $\underline{\underline{A}}$, that form a complete basis. Indeed supposing

$$\underline{\underline{A}} = \sum_{\sigma} \underline{v}_{\sigma} a_{\sigma} \underline{v}_{\sigma}^{\dagger}, \qquad (135)$$

with

$$\delta_{\sigma\sigma'} = \underline{v}_{\sigma}^{\dagger}\underline{v}_{\sigma'}, \qquad (136)$$

one can find the expansion coefficients of $\underline{y}$,

$$\underline{y} = \sum_{\sigma} y_{\sigma}\underline{v}_{\sigma}, \qquad (137)$$

then build $\underline{x}$ as

$$\underline{x} = \sum_{\sigma} \frac{y_{\sigma}}{a_{\sigma} - \epsilon}\underline{v}_{\sigma}, \qquad (138)$$

where ϵ is determined by the constraint $\underline{x}^{\dagger}\underline{x} = 1$, namely

$$\sum_{\sigma} \frac{y_{\sigma}^{\dagger}y_{\sigma}}{(a_{\sigma} - \epsilon)^2} = 1 \qquad (139)$$

We can characterize the set of allowed ϵ as follows. In the case the inhomogeneous term is suppressed in Eq. (134), every eigenvalue a_{σ} belongs to this set. Then, supposing at least y_{σ} is non-vanishing, the function in the left-hand side of Eq. (134) is always positive, and made of separate positive contributions, each diverging positively at their corresponding a_{σ}. In each interval separating two consecutive a_{σ} and $a_{\sigma'}$, the function first decreases monotonically, reaches a positive minimum, then increases monotonically. Its second derivative is always positive. Thus, between two such consecutive eigenvalues, there are two, one or zero solution of Eq. 139. For small values of y_{σ}, each initial solution $\epsilon = a_{\sigma}$ will double, with one solution slightly below a_{σ} and one solution above a_{σ}. With increasing scaling of y_{σ}, the left-hand side function will become higher than 1 in different intervals between eigenvalues a_{σ}, and only the solution with ϵ much lower than all a_{σ}, or much larger than all a_{σ} will survive. There is always one unique solution below the lowest a_{σ}, and one unique solution above the highest a_{σ}.

B.2 Stability analysis

Despite this proof of the existence of one low and one high solution, the stability of the solutions of this equation with respect to variations of $\underline{y}$ might be problematic, including when $\underline{y}$ is small. This might seem counter-intuitive, as we are then close to a usual eigenvalue problem without inhomogeneous term.

Indeed, suppose the set of inhomogeneous y_{σ} is perturbed by δy_{σ}. Then, at first order, one deduces from Eq. 138:

$$\delta x_{\sigma} = \frac{\delta y_{\sigma}}{a_{\sigma} - (\epsilon + \delta\epsilon)}, \qquad (140)$$

with

$$\delta\epsilon = \sum_{\sigma} -\frac{1}{2}(a_{\sigma} - \epsilon)\big(\delta y_{\sigma}^{\dagger}y_{\sigma} + y_{\sigma}^{\dagger}\delta y_{\sigma}\big). \qquad (141)$$

As discussed previously, when $\underline{y}$ is small, ϵ is close to one of the a_{σ}. Hence, the possibly small variation of the corresponding δy_{σ} will be amplified by the inverse of $(a_{\sigma} - \epsilon)$ in Eq. 140, which might be very big.

Moreover, such an analysis reveals the difficulty to treat a coupled system of equations such as

$$\underline{\underline{A}}_1\underline{x} + \underline{y} = \epsilon_1\underline{x}$$
$$\underline{\underline{A}}_2\underline{y} + \underline{x} = \epsilon_2\underline{y} \qquad (142)$$

with $\underline{x}^{\dagger}\underline{x} = 1$ and $\underline{y}^{\dagger}\underline{y} = 1$, by separately treating them sequentially, one after the other: the consecutive feedbacks might amplify and lead to divergence.

References

1. A. Fetter, J. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 2000)
2. G.D. Mahan, *Many-Particle Physics* (Kluwer, New York, 2000)
3. R. Martin, L. Reining, D. Ceperley, *Interacting Electrons. Theory and Computational Approaches* (Cambrige University Press, Cambridge, 2016)
4. M. Di Ventra, *Non-equilibrium Green's function formalism.*, in *Electrical Transport in Nanoscale Systems* (Cambrige University Press, 2008), p. 209
5. G. Kotliar, S. Savrasov, K. Haule, V. Oudovenko, O. Parcollet, C. Marianetti, Rev. Mod. Phys. **78**, 865 (2006)
6. T. Azumi, K. Matsuzaki, Photochemistry and photobiology **25**, 315 (1977)
7. R. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, 2008)
8. M. Born, K. Huang, *The dynamical theory of crystal lattices* (Oxford University Press (London), 1954)
9. G. Hunter, International Journal of Quantum Chemistry **9**, 237 (1975)
10. N.I. Gidopoulos, E. Gross, arXiv: cond-mat.mtrl-sci **0502**, 433 (2005)
11. N.I. Gidopoulos, E.K.U. Gross, Phil. Trans. R. Soc. **372**, 0059 (2014)
12. A. Abedi, N.T. Maitra, E.K.U. Gross, Phys. Rev. Lett. **105**, 123002 (2010)
13. A. Abedi, N.T. Maitra, E.K.U. Gross, J. Chem. Phys. **137**, 22A530 (2012)
14. M.V. Berry, Proc. R. Soc. Lond. A **392**, 45 (1984)
15. Y. Suzuki, A. Abedi, N.T. Maitra, K. Yamashita, E.K.U. Gross, Phys. Rev. A **89**, 040501 (2014)
16. L.S. Cederbaum, Chem. Phys. **457**, 129 (2015)
17. G. Hunter, International Journal of Quantum Chemistry **29**, 197 (1986)
18. A. Schild, E.K.U. Gross, Phys. Rev. Lett. **118**, 163202 (2017)
19. D. Bohm, Phys. Rev. **85**, 166 (1952)
20. A. Abedi, F. Agostini, E.K.U. Gross, Europhys. Lett. **106**, 33001 (2014)
21. F. Agostini, A. Abedi, E. Gross, J. Chem. Phys. **141**, 214101 (2014)
22. S.K. Min, F. Agostini, E.K.U. Gross, Phys. Rev. Lett. **115**, 073001 (2015)
23. F. Agostini, S.K. Min, A. Abedi, E.K.U. Gross, J. Chem. Theory Comput. **12**, 2127 (2016)
24. S.K. Min, F. Agostini, I. Tavernelli, E.K.U. Gross, J. Phys. Chem. Lett. **8**, 3048 (2017)
25. R. Requist, E.K.U. Gross, Phys. Rev. Lett. **117**, 193001 (2016)
26. R. Requist, C.R. Proetto, E.K.U. Gross, Phys. Rev. A **96**, 062503 (2017)
27. J. Tully, J. Chem. Phys. **93**, 1061 (1990)
28. G. Granucci, M. Persico, J. Chem. Phys. **126**, 134114 (2007)
29. P. Ehrenfest, Zeitschrift für Physik **45**, 455 (1927)
30. E. Khosravi, A. Abedi, N.T. Maitra, Phys. Rev. Lett. **115**, 263002 (2015)
31. E. Khosravi, A. Abedi, A. Rubio, N.T. Maitra, Phys. Chem. Chem. Phys. **19**, 8269 (2017)
32. Y. Suzuki, A. Abedi, N.T. Maitra, E.K.U. Gross, Phys. Chem. Chem. Phys. **17**, 29271 (2015)
33. T. Fiedlschuster, J. Handt, E.K.U. Gross, R. Schmidt, Phys. Rev. A **95**, 063424 (2017)
34. G. Albareda, A. Abedi, I. Tavernelli, A. Rubio, Phys. Rev. A **94**, 062511 (2016)
35. T. Holstein, Annals of physics **8**, 325 (1959)
36. T. Holstein, Annals of physics **8**, 343 (1959)
37. R.L. Fulton, M. Gouterman, J. Chem. Phys. **35**, 1059 (1961)
38. L. Zijing, H. Rongsheng, W. Kelin, International Journal of Modern Physics B **17**, 4252 (2003)
39. L.K. McKemmish, R.H. McKenzie, N.S. Hush, J.R. Reimers, J. Chem. Phys. **135**, 244110 (2011)
40. D. Braak, Phys. Rev. Lett. **107**, 100401 (2011)
41. I.B. Bersuker, Chem. Rev. **113**, 1351 (2013)
42. N. Säkkinen, Y. Peng, H. Appel, R. van Leeuwen, J. Chem. Phys. **143**, 234101 (2015)
43. E.R.J. Vandaele, A. Arvanitidis, A. Ceulemans, J. Phys. A: Math. Theor **50**, 114002 (2017)
44. H. van der Vorst, *Iterative Krylov Methods for Large Linear Systems* (Cambridge University Press, 2009)
45. M.R. Hestenes, E. Stiefel, J. Res. Nat. Bur. Stand **49**, 409 (1952)
46. W. Press, B. Flannery, S. Teukolsky, W. Vetterling, *Numerical Recipes. The art of scientific programming* (Cambridge University Press, Cambridge, 1988)
47. M. Teter, M. Payne, D. Allan, Phys. Rev. B **40**, 12255 (1989)
48. M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, J.D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992)
49. I.G. Lang, Y.A. Firsov, Sov. Phys. JETP **16**, 1301 (1963)
50. D. Langreth, Phys. Rev. B **1**, 471 (1970)
51. F. Agostini, A. Abedi, Y. Suzuki, S. Min, N. Maitra, E. Gross, J. Chem. Phys. **142**, 084303 (2015)