

# A biorthonormal formalism for nonadiabatic coupled cluster dynamics

Eirik F. Kjørstad<sup>†</sup> and Henrik Koch<sup>\*,‡,†</sup>

<sup>†</sup>*Department of Chemistry, Norwegian University of Science and Technology, 7491  
Trondheim, Norway*

<sup>‡</sup>*Scuola normale superiore, Piazza dei Cavalieri, 7, 56126 Pisa PI, Italy*

E-mail: henrik.koch@sns.it

## Abstract

In coupled cluster methods, the electronic states are biorthonormal in the sense that the left states are orthonormal to the right states. Here we present an extension of this formalism to a left and right total molecular wave function. Starting from left and right Born-Huang expansions, we derive projected Schrödinger equations for the left and right nuclear wave functions. Observables may be extracted from the resulting wave function pair using standard expressions. The formalism is shown to be invariant under electronic basis transformations, such as normalization of the electronic states. Consequently, the nonadiabatic coupling elements can be expressed with biorthonormal wave functions. Calculating normalization factors that scale as full-CI is therefore not necessary, contrary to claims in the literature. For nuclear dynamics, we therefore need expressions for the vector and scalar couplings in the biorthonormal formalism. We derive these expressions using a Lagrangian formalism.

# Introduction

Nonadiabatic coupling elements account for electron-nucleus interactions that are neglected in the Born-Oppenheimer<sup>1</sup> (BO) approximation. These elements couple different electronic states through the nuclear kinetic energy operator. While mostly negligible in ground state chemistry, coupling elements are required when considering molecular dynamics in excited electronic states. Excited state dynamics often involves regions of nuclear space where electronic states are nearly or exactly degenerate, causing a breakdown of the BO separation.<sup>2,3</sup> Accurately describing nonadiabatic coupling elements is therefore important for reliable predictions in photochemistry.

The coupled cluster method is one of the most accurate electronic structure methods, both for ground and excited state properties,<sup>4-7</sup> but it has not found widespread use for predicting excited state dynamics. This is primarily because standard coupled cluster methods give a nonphysical description of regions close to electronic degeneracies, or conical intersections.<sup>8-10</sup> This issue can be traced to the method’s non-Hermiticity, which seems to imply that coupled cluster methods cannot be used for nonadiabatic dynamics. However, this is not the case. As we have shown in recent work, the method can be constrained to give a correct physical description of excited state conical intersections while retaining the standard non-Hermitian formalism and presumably its accuracy.<sup>11,12</sup> These developments may lead to renewed interest in coupled cluster dynamics.

Nonadiabaticity, as described by coupled cluster methods, has been considered by several authors. The formula for the vector coupling was first derived by Christiansen,<sup>13</sup> who applied the  $Z$ -vector substitution method<sup>14</sup> on a biorthonormal expression for the vector coupling,

$$\mathbf{F}_{mn}^I = \langle \tilde{\psi}_m | \nabla_I \psi_n \rangle, \quad \langle \tilde{\psi}_m | \psi_n \rangle = \delta_{mn}, \quad (1)$$

where  $(\tilde{\psi}_k, \psi_k)$  refers to the left and right  $k$ th electronic states, and  $I$  to a nucleus. However, Christiansen’s paper<sup>13</sup> did not include an implementation of the coupling. The vector cou-

pling was later rederived by Tajti and Szalay<sup>15</sup> by differentiating the corresponding  $m$ -to- $n$  transition element of the electronic Hamiltonian. Their derivation is closely related to that given by Ichino *et al.*<sup>16</sup> for the quasidiabatic interstate coupling. Tajti and Szalay<sup>15</sup> also gave an implementation of the vector coupling at the singles and doubles level (CCSD<sup>17</sup>). These papers on the vector coupling<sup>13,15</sup> did not include a discussion of the nuclear Schrödinger equations in coupled cluster theory, where the coupling elements enter.

The correct formula for the vector coupling has been a subject of some controversy. Tajti and Szalay<sup>15</sup> argued that the biorthonormal formula in Eq. (1) is incorrect. As they correctly noted, the vector coupling changes with the norm of the left and right states. A similar observation had been made in an earlier paper on the diagonal BO correction.<sup>18</sup> Since the vector coupling varies with the norm of the states, the full-CC vector coupling is different from the full configuration interaction (CI) limit, where left and right states are identical and usually normalized. They therefore suggested that normalizing the states was necessary. Furthermore, since the derivative can either act on the left or on the right state, they suggested using an average of the two.<sup>15</sup> If true, these observations are troubling: they imply that computing the vector coupling has a computational cost that scales as full-CI due to the normalization factors for the right states. In practice, the normalization factors are therefore approximated. However, it is unfortunate if one must resort to approximations other than the truncation level of the coupled cluster method (e.g., singles and doubles). The need for normalization factors was also assumed in the recent CCSD implementation by Faraji *et al.*<sup>19</sup>

One of the main objectives of the present paper is to establish that normalization is not necessary. The reason is that normalization is a special case of an invertible transformation of the electronic basis. Such transformations do not change the expansion space in the Born-Huang expansion<sup>20</sup> and therefore do not change the molecular wave function. In particular, the coefficients in the Born-Huang expansion—that is, the nuclear wave functions—absorb the transformation of the electronic states. In a recent paper, Shamasundar<sup>21</sup> also noted that

the predicted dynamics must not depend on the normalization of the underlying electronic wave functions. The vector coupling does depend on normalization, but this should not be considered a problem because this quantity is not an observable. Since normalization of wave functions is not necessary, the biorthonormal formula in Eq. (1) is a valid option. In this work, we derive the biorthonormal coupled cluster vector and scalar couplings using the Lagrangian approach developed by Hohenstein.<sup>22</sup>

The second main objective of the paper is to give a framework for nonadiabatic dynamics using coupled cluster methods. In particular, we argue that the biorthonormal formalism for electronic wave functions implies a biorthonormal formalism for the molecular wave function. Hence, we must determine left and right nuclear wave functions and the nuclear motion is described by two sets of nuclear Schrödinger equations. The result is a molecular wave function pair  $(\tilde{\Psi}, \Psi)$ , where observable quantities are given by the usual biorthonormal formulas. In this contribution, we describe theoretical aspects relevant for nonadiabatic dynamics. Implementation of the various quantities is postponed to a future publication.

## Theory

The total wave function of a molecular system can be expressed as an expansion over the electronic wave functions. The coefficients of this Born-Huang expansion defines the nuclear wave functions. These are determined by inserting the expansion in the Schrödinger equation and projecting out the electronic components. To formulate the corresponding procedure for coupled cluster theory, we first review the description of the electronic states.

## Electronic wave functions in coupled cluster theory

In the equation of motion coupled cluster formalism, a set of left and right electronic states are considered. These are defined as<sup>23</sup>

$$\langle \tilde{\psi}_n | = \sum_{\mu \geq 0} \mathcal{L}_\mu^n \langle \mu | \exp(-T), \quad (2)$$

$$|\psi_n\rangle = \sum_{\mu \geq 0} \exp(T) |\mu\rangle \mathcal{R}_\mu^n, \quad n = 0, 1, 2, \dots \quad (3)$$

The states are not identical, in general, but they satisfy the biorthonormality condition

$$\langle \tilde{\psi}_m | \psi_n \rangle = \delta_{mn}. \quad (4)$$

The scalars  $\mathcal{L}_\mu^n$  and  $\mathcal{R}_\mu^n$  are state amplitudes, giving weights to the configurations

$$|\mu\rangle = \tau_\mu |\text{HF}\rangle \quad (5)$$

$$\langle \mu | = \langle \text{HF} | \tilde{\tau}_\mu^\dagger, \quad \mu \geq 0, \quad (6)$$

where  $\tau_\mu$  and  $\tilde{\tau}_\mu$  with  $\mu > 0$  are excitation operators relative to the Hartree-Fock state  $|\text{HF}\rangle$ , while  $\tau_0 = \tilde{\tau}_0 = \mathbb{I}$  is the identity operator. The ket and bra bases,  $\{|\mu\rangle\}$  and  $\{\langle \mu | \}$ , span the same subspace and are normally required to satisfy the biorthonormality relation

$$\langle \mu | \nu \rangle = \delta_{\mu\nu}, \quad \mu, \nu \geq 0. \quad (7)$$

One special case is that the left and right bases are identical and hence orthonormal. However, using different left and right basis is sometimes convenient (e.g., in spin-adapted formulations<sup>6</sup>). Finally, we have the exponential part of the parametrization, defined by the

cluster operator

$$T = \sum_{\mu>0} t_{\mu} \tau_{\mu}. \quad (8)$$

The scalars  $t_{\mu}$  are called cluster amplitudes.

Given the parametrization in Eqs. (2) and (3), how are the parameters determined? First one assumes that the right ground state can be written

$$|\psi_0\rangle = \exp(T)|\text{HF}\rangle. \quad (9)$$

Then the time-independent Schrödinger equation, expressed as

$$\exp(-T)H \exp(T)|\text{HF}\rangle = E_0|\text{HF}\rangle, \quad (10)$$

is projected onto the bra basis. The operator

$$\bar{H} = \exp(-T)H \exp(T) \quad (11)$$

is known as the similarity transformed Hamiltonian. This projection procedure gives an expression for the ground state energy and equations for determining the amplitudes,

$$E_0 = \langle \text{HF} | \bar{H} | \text{HF} \rangle \quad (12)$$

$$\Omega_{\mu} = \langle \mu | \bar{H} | \text{HF} \rangle = 0, \quad \mu > 0. \quad (13)$$

The state amplitudes are determined by making the pseudo expectation values

$$E_n(\mathcal{L}_n, \mathcal{R}_n) = \langle \tilde{\psi}_n | \bar{H} | \psi_n \rangle, \quad n = 0, 1, 2, \dots \quad (14)$$

stationary under the binormality condition given in Eq. (4). This constrained optimization

problem is conveniently formulated in terms of the state Lagrangians

$$\begin{aligned}\mathcal{L}_n(\mathcal{L}_n, \mathcal{R}_n, \bar{E}_n) &= \langle \tilde{\psi}_n | \bar{H} | \psi_n \rangle + \bar{E}_n(1 - \langle \tilde{\psi}_n | \psi_n \rangle) \\ &= \mathcal{L}_n^T \bar{\mathcal{H}} \mathcal{R}_n + \bar{E}_n(1 - \mathcal{L}_n^T \mathcal{R}_n),\end{aligned}\tag{15}$$

where we have defined the Hamiltonian matrix

$$\bar{\mathcal{H}}_{\mu\nu} = \langle \mu | \bar{H} | \nu \rangle, \quad \mu, \nu \geq 0.\tag{16}$$

Stationarity of the Lagrangians imply

$$1 = \mathcal{L}_n^T \mathcal{R}_n\tag{17}$$

$$\mathbf{0} = \mathcal{L}_n^T \bar{\mathcal{H}} - \bar{E}_n \mathcal{L}_n^T\tag{18}$$

$$\mathbf{0} = \bar{\mathcal{H}} \mathcal{R}_n - \bar{E}_n \mathcal{R}_n.\tag{19}$$

Writing  $\bar{E}_n = E_n = E_0 + \omega_n$ , we see that the latter two equations read

$$\mathcal{L}_n^T \mathcal{A} = \omega_n \mathcal{L}_n^T\tag{20}$$

$$\mathcal{A} \mathcal{R}_n = \omega_n \mathcal{R}_n, \quad n = 0, 1, 2, \dots,\tag{21}$$

where  $\mathcal{A} = \bar{\mathcal{H}} - E_0 \mathbf{I}$ . Identifying the pseudo expectation value in Eq. (14) with the energy, we see that  $\omega_n$  is the excitation energy of the  $n$ th state, where it is understood that  $\omega_0 = 0$  for the ground state. The matrix  $\mathcal{A}$  can be expressed as

$$\mathcal{A} = \begin{pmatrix} 0 & \boldsymbol{\eta}^T \\ \mathbf{0} & \mathbf{A} \end{pmatrix}\tag{22}$$

where

$$\eta_\nu = \langle \text{HF} | [\bar{H}, \tau_\nu] | \text{HF} \rangle \quad (23)$$

$$A_{\mu\nu} = \langle \tilde{\mu} | [\bar{H}, \tau_\nu] | \text{HF} \rangle. \quad (24)$$

where  $\mathbf{A}$  is called the coupled cluster Jacobian matrix.<sup>24</sup> The eigenvalues of  $\mathbf{A}$  are the non-zero excitation energies, i.e.  $\omega_n$  with  $n = 1, 2, 3, \dots$

## The Born-Huang expansion of the total wave function and the nuclear Schrödinger equations

With the electronic states described, we now turn to the expansion of the total wave function. The Born-Huang expansion expresses the total wave function in terms of the left and right electronic bases given in Eqs. (2) and (3). Notice that this implies a biorthonormal description of the total wave function, since we can expand in both the left and right states. Hence we have a left and a right total wave function

$$\Psi(\mathbf{r}, \mathbf{R}, t) = \sum_n \chi_n(\mathbf{R}, t) \psi_n(\mathbf{r}; \mathbf{R}) \quad (25)$$

$$\tilde{\Psi}(\mathbf{r}, \mathbf{R}, t) = \sum_n \tilde{\chi}_n(\mathbf{R}, t) \tilde{\psi}_n(\mathbf{r}; \mathbf{R}), \quad (26)$$

with associated left and right nuclear wave functions  $\tilde{\chi}_n$  and  $\chi_n$ , and

$$1 = \langle \tilde{\Psi} | \Psi \rangle = \sum_{mn} \langle \tilde{\chi}_m | \langle \tilde{\psi}_m | \psi_n \rangle | \chi_n \rangle = \sum_{mn} \delta_{mn} \langle \tilde{\chi}_m | \chi_n \rangle = \sum_n \langle \tilde{\chi}_n | \chi_n \rangle, \quad (27)$$

where we have assumed biorthonormal electronic states in the third equality. Expectation values are defined through the standard expression<sup>23,24</sup>

$$\langle \Omega \rangle = \langle \tilde{\Psi} | \Omega | \Psi \rangle, \quad \Omega = \Omega^\dagger. \quad (28)$$



To derive the equations for the nuclear wave functions, one normally projects the total Schrödinger equation on the electronic basis. In this respect, a biorthonormal description is advantageous; for practical coupled cluster models, where the excitation space is truncated to some excitation order, projection of the *right* Schrödinger equation is done onto the *left* electronic basis, leading to computationally tractable expressions that scale as expected for the given model (e.g.  $\mathcal{O}(N^6)$  for CCSD).

By inserting the  $\Psi$  in Eq. (25) into the time dependent Schrödinger equation,

$$H\Psi = i\frac{d\Psi}{dt}, \quad (29)$$

and projecting it onto the left electronic basis, we get a coupled set of equations for the right nuclear wave functions  $\chi_n$ . These nuclear Schrödinger equations can be expressed as

$$(i\frac{d}{dt} - E_m)\chi_m = \sum_{I,n} \frac{1}{2M_I} (\delta_{mn}\nabla_I^2 + G_{mn}^I + 2\mathbf{F}_{mn}^I \cdot \nabla_I)\chi_n, \quad (30)$$

where we have suppressed the  $\mathbf{R}$  and  $t$  dependence for readability. The nonadiabatic coupling vectors in Eq. (30) are given in the biorthonormal basis:

$$G_{mn}^I = \langle \tilde{\psi}_m | \nabla_I^2 \psi_n \rangle \quad (31)$$

$$\mathbf{F}_{mn}^I = \langle \tilde{\psi}_m | \nabla_I \psi_n \rangle. \quad (32)$$

These are called the scalar and vector couplings, respectively.

In analogous fashion, we derive the nuclear Schrödinger equations for the left nuclear wave functions from the complex conjugated Schrödinger equation

$$H\tilde{\Psi} = -i\frac{d\tilde{\Psi}}{dt}. \quad (33)$$

Inserting Eq. (26) into Eq. (33), and projecting onto the left electronic basis, leads to

$$\left(-i\frac{d}{dt} - E_m\right)\tilde{\chi}_m = \sum_{I,n} \frac{1}{2M_I} (\delta_{nm}\nabla_I^2 + \tilde{G}_{nm}^I + 2\tilde{\mathbf{F}}_{nm}^I \cdot \nabla_I)\tilde{\chi}_n, \quad (34)$$

where

$$\tilde{G}_{nm}^I = \langle \nabla_I^2 \tilde{\psi}_n | \psi_m \rangle \quad (35)$$

$$\tilde{\mathbf{F}}_{nm}^I = \langle \nabla_I \tilde{\psi}_n | \psi_m \rangle. \quad (36)$$

The nuclear Schrödinger equations may be expressed in the more compact matrix notation

$$\left(i\frac{d}{dt} - \mathbf{E}\right)\boldsymbol{\chi} = \sum_I \frac{1}{2M_I} (\mathbf{I}\nabla_I^2 + \mathbf{G}_I + 2\mathbf{F}_I \cdot \nabla_I)\boldsymbol{\chi} \quad (37)$$

$$\left(-i\frac{d}{dt} - \mathbf{E}\right)\tilde{\boldsymbol{\chi}} = \sum_I \frac{1}{2M_I} (\mathbf{I}\nabla_I^2 + \tilde{\mathbf{G}}_I + 2\tilde{\mathbf{F}}_I \cdot \nabla_I)\tilde{\boldsymbol{\chi}}, \quad (38)$$

where  $\mathbf{E}$  is a diagonal matrix with the electronic energies on the diagonal,  $\mathbf{I}$  is the identity matrix,  $\boldsymbol{\chi}$  is a vector containing the right nuclear wave functions, and  $\mathbf{G}_I$  and  $\mathbf{F}_I$  are matrices consisting of the scalar and vector couplings of the  $I$ th nucleus, respectively. The quantities with a tilde are similarly defined.

This matrix notation has been used to illuminate some relations to gauge theories in the nuclear Schrödinger equations; Pacher *et al.*<sup>25</sup> found that the vector coupling can be seen to serve a role analogous to the vector potential in electromagnetism. In the present work, it serves as a useful notation for dealing with basis transformations and the vector algebra needed to demonstrate invariance under such transformations.

## Basis invariance and the special case of norm invariance

In the literature on nonadiabatic coupling vectors in coupled cluster theory, normalization is often considered problematic. The reason is that the left and right states are binormal in the

coupled cluster formalism. Compared to the nonadiabatic couplings in full-CI theory, where the states are normalized, the full coupled cluster limit is “incorrect” because the value of the couplings depend on the geometry-dependent normalization constants. While this suggests that one should normalize the states, doing so is not straightforward. The computational cost of the normalization factor scales as full-CI for the right electronic states:<sup>16</sup>

$$N_R^n = \langle \psi_n | \psi_n \rangle \quad (39)$$

$$N_L^n = \langle \tilde{\psi}_n | \tilde{\psi}_n \rangle. \quad (40)$$

Since one cannot evaluate  $N_R^n$  in general, some have suggested  $N_R^n = (N_L^n)^{-1}$  or  $N_R^n = N_L^n = 1$  as alternatives. The former gives the full-CI limit while the latter simply assumes the standard binormality.<sup>16,19</sup>

Binormality is not an issue from the point of view of dynamics. Changing the norm of the electronic states is a special case of a basis transformation of the electronic basis. As such, the Born-Huang expansion and the projection equations are equivalent in the transformed and untransformed bases. Changes in the electronic basis are absorbed in the expansion coefficients, i.e., the nuclear wave functions. In the special case of normalization, the right electronic wave functions are divided by  $N_R^n$  while the right nuclear wave functions are multiplied by  $N_R^n$ . The total wave function is invariant under such transformations.

More precisely, consider invertible transformations of the left and right electronic bases. In vector notation, these transformations can be expressed as

$$\tilde{\psi}' = \tilde{\psi} \mathbf{N} \quad (41)$$

$$\psi' = \psi \mathbf{M}, \quad (42)$$

where the matrices  $\mathbf{M}$  and  $\mathbf{N}$  are assumed to be smooth invertible matrix functions of the nuclear coordinates. For notational simplicity, we have let the left and right wave function vectors be row vectors. Transformed quantities are denoted by a prime. In the transformed

basis, the total left and right wave function have the Born-Huang expansions

$$\Psi'(\mathbf{r}, \mathbf{R}, t) = \sum_n \chi'_n(\mathbf{R}, t) \psi'_n(\mathbf{r}; \mathbf{R}) \quad (43)$$

$$\tilde{\Psi}'(\mathbf{r}, \mathbf{R}, t) = \sum_n \tilde{\chi}'_n(\mathbf{R}, t) \tilde{\psi}'_n(\mathbf{r}; \mathbf{R}). \quad (44)$$

We wish to show that the wave function in the transformed basis is identical to that obtained in the untransformed basis; that is,  $\Psi' = \Psi$  and  $\tilde{\Psi}' = \tilde{\Psi}$ . The conclusion that follows is that the choice of electronic basis does not change the predictions of the theory. In other words, it is perfectly appropriate to use the biorthonormal description that is standard in coupled cluster theory.<sup>24</sup>

Before proceeding, we define some notation. In the transformed basis, we have to account for the non-unit overlap of the electronic wave functions. Hence, when projecting the time-dependent Schrödinger equation onto the electronic basis, we get electronic overlap matrix elements. We define these elements as

$$S_{mn} = \langle \tilde{\psi}'_m | \psi'_n \rangle = \sum_{kl} \langle \tilde{\psi}_k N_{km} | \psi_l M_{ln} \rangle = \sum_{kl} N_{km}^* \delta_{kl} M_{ln} = (\mathbf{N}^\dagger \mathbf{M})_{mn}. \quad (45)$$

Similarly, the electronic Hamiltonian matrix is not necessarily diagonal:

$$(\mathbf{H}_e)_{mn} = \langle \tilde{\psi}'_m | H_e \psi'_n \rangle = \sum_{kl} \langle \tilde{\psi}_k N_{km} | H_e \psi_l M_{ln} \rangle = \sum_{kl} N_{km}^* E_{kl} M_{ln} = (\mathbf{N}^\dagger \mathbf{E} \mathbf{M})_{mn}. \quad (46)$$

We show the equivalence for the right wave functions. The proof for the left wave function is identical. Following the standard procedure, we now insert the transformed wave function in Eq. (43) into the Schrödinger equation and project onto the transformed left electronic wave functions. The result is the right nuclear Schrödinger equation

$$(\mathbf{N}^\dagger \mathbf{E} \mathbf{M} - i \frac{d}{dt} \mathbf{S}) \boldsymbol{\chi}' = \sum_I \frac{1}{2M_I} (\mathbf{S} \nabla_I^2 + \mathbf{G}'_I + 2\mathbf{F}'_I \cdot \nabla_I) \boldsymbol{\chi}'. \quad (47)$$

If the total wave function is invariant, and  $\psi' = \psi \mathbf{M}$ , then we must have nuclear wave functions that cancel the transformation of the electronic wave functions:

$$\chi' = \mathbf{M}^{-1} \chi. \quad (48)$$

Indeed, with  $\chi'$  as given in Eq. (48), we have

$$\Psi' = \sum_k \psi'_k \chi'_k = \sum_{klm} \psi_l M_{lk} M_{km}^{-1} \chi_m = \sum_l \psi_l \chi_l = \Psi. \quad (49)$$

Let us confirm that Eq. (48) is in fact a solution to the transformed nuclear Schrödinger equation in Eq. (47). We begin by relating the old and new nonadiabatic coupling terms. The gradient of the electronic wave functions transform as

$$\nabla_I \psi'_l = \sum_m \nabla_I (\psi_m M_{ml}) = \sum_m \left( (\nabla_I \psi_m) M_{ml} + \psi_m (\nabla_I M_{ml}) \right). \quad (50)$$

Hence, the vector couplings can be written as

$$(\mathbf{F}'_I)_{kl} = \langle \tilde{\psi}'_k | \nabla_I \psi'_l \rangle = \sum_n N_{nk}^* \langle \tilde{\psi}_n | \nabla_I \psi'_l \rangle = \sum_{nm} \left( N_{nk}^* F_{nm}^I M_{ml} + N_{nk}^* \delta_{nm} \nabla_I M_{ml} \right). \quad (51)$$

In more compact matrix notation, we have

$$\mathbf{F}'_I = \mathbf{N}^\dagger \mathbf{F}_I \mathbf{M} + \mathbf{N}^\dagger (\nabla_I \mathbf{M}). \quad (52)$$

Similarly, the Laplacian of the electronic wave functions transform as

$$\nabla_I^2 \psi'_l = \nabla_I \cdot \nabla_I \psi'_l = \sum_m \left( (\nabla_I^2 \psi_m) M_{ml} + 2(\nabla_I \psi_m) \cdot (\nabla_I M_{ml}) + \psi_m (\nabla_I^2 M_{ml}) \right), \quad (53)$$

implying that the scalar couplings transform as

$$\mathbf{G}'_I = \mathbf{N}^\dagger \mathbf{G}_I \mathbf{M} + 2\mathbf{N}^\dagger \mathbf{F}_I \cdot (\nabla_I \mathbf{M}) + \mathbf{N}^\dagger (\nabla_I^2 \mathbf{M}). \quad (54)$$

The gradient and Laplacian of  $\chi'$  is derived in the same way as for the electronic states, giving

$$\nabla_I \chi'_l = \sum_m (M_{lm}^{-1} (\nabla_I \chi_m) + (\nabla_I M_{lm}^{-1}) \chi_m) \quad (55)$$

$$\nabla_I^2 \chi'_l = \sum_m (M_{lm}^{-1} (\nabla_I^2 \chi_m) + 2(\nabla_I M_{lm}^{-1}) \cdot (\nabla_I \chi_m) + (\nabla_I^2 M_{lm}^{-1}) \chi_m). \quad (56)$$

Thus, we have the following contributions on the right hand side of the nuclear Schrödinger equation:

$$\mathbf{S} \nabla_I^2 \chi' = \mathbf{N}^\dagger (\nabla_I^2 \chi + 2\mathbf{M} (\nabla_I \mathbf{M}^{-1}) \cdot \nabla_I \chi + \mathbf{M} (\nabla_I^2 \mathbf{M}^{-1}) \chi) \quad (57)$$

$$\mathbf{G}'_I \chi' = \mathbf{N}^\dagger (\mathbf{G}_I \chi + 2\mathbf{F}_I \cdot (\nabla_I \mathbf{M}) \mathbf{M}^{-1} \chi + (\nabla_I^2 \mathbf{M}) \mathbf{M}^{-1} \chi) \quad (58)$$

$$2\mathbf{F}'_I \cdot \nabla_I \chi' = \mathbf{N}^\dagger (2\mathbf{F}_I \cdot \nabla_I \chi + 2\mathbf{F}_I \cdot \mathbf{M} (\nabla_I \mathbf{M}^{-1}) \chi + 2(\nabla_I \mathbf{M}) \mathbf{M}^{-1} \cdot \nabla_I \chi + 2(\nabla_I \mathbf{M}) \cdot (\nabla_I \mathbf{M}^{-1}) \chi). \quad (59)$$

Though somewhat involved, most of the terms cancel when added together. In fact, since

$$0 = \nabla_I (\mathbf{M} \mathbf{M}^{-1}) = (\nabla_I \mathbf{M}) \mathbf{M}^{-1} + \mathbf{M} (\nabla_I \mathbf{M}^{-1}) \quad (60)$$

$$0 = \nabla_I^2 (\mathbf{M} \mathbf{M}^{-1}) = (\nabla_I^2 \mathbf{M}) \mathbf{M}^{-1} + \mathbf{M} (\nabla_I^2 \mathbf{M}^{-1}) + 2(\nabla_I \mathbf{M}) \cdot (\nabla_I \mathbf{M}^{-1}), \quad (61)$$

we can write

$$\begin{aligned}
S\nabla_I^2\chi' + \mathbf{G}'_I\chi' + 2\mathbf{F}'_I \cdot \nabla_I\chi' &= \mathbf{N}^\dagger(\nabla_I^2\chi + \mathbf{G}_I\chi + 2\mathbf{F}_I \cdot \nabla_I\chi \\
&+ 2\nabla_I(\mathbf{M}\mathbf{M}^{-1}) \cdot \nabla_I\chi + \nabla_I^2(\mathbf{M}\mathbf{M}^{-1})\chi \\
&+ 2\mathbf{F}_I \cdot \nabla_I(\mathbf{M}\mathbf{M}^{-1})\chi) \\
&= \mathbf{N}^\dagger(\nabla_I^2\chi + \mathbf{G}_I\chi + 2\mathbf{F}_I \cdot \nabla_I\chi).
\end{aligned} \tag{62}$$

In other words, with  $\chi' = \mathbf{M}^{-1}\chi$ , the right nuclear Schrödinger equation simplifies to

$$(\mathbf{N}^\dagger \mathbf{E} - i\frac{d}{dt}\mathbf{N}^\dagger)\chi = \mathbf{N}^\dagger \sum_I \frac{1}{2M_I}(\nabla_I^2 + \mathbf{G}_I + 2\mathbf{F}_I \cdot \nabla_I)\chi, \tag{63}$$

which, upon premultiplication by  $\mathbf{N}^{-\dagger}$ , is seen to be equivalent to the original right nuclear Schrödinger equation in Eq. (37).

Since all the derivation steps we have made are reversible, we have shown that  $\chi$  is a solution to the untransformed nuclear Schrödinger equation if and only if  $\chi'$  is a solution to the transformed Schrödinger equation. The total right wave function is therefore invariant with respect to transformations of the electronic basis,  $\Psi' = \Psi$ .

One consequence of basis invariance is that the nonadiabatic couplings can be derived in the standard biorthonormal formalism. To derive expressions for these elements, we must first consider the geometry dependence of the many-body operators.

## Geometry dependence of the many-body operators

The scalar and vectors couplings, see Eqs. (154) and (32), involve differentiation of the electronic wave functions with respect to the nuclear coordinates  $\mathbf{x}$ . To evaluate these, we need to consider the dependence of both the wave function parameters and the many-body operators. The operator's dependence is handled through orbital connections which relates orbitals at neighbouring geometries. Note that there is no unique orbital connection; many-body operators are expressed with respect to a specific orthonormal orbital basis, but at each

geometry there are an infinite number of such bases related by unitary transformations. For reasons that will become clear, we will use the so-called natural connection. Our presentation will follow closely that given by of Olsen *et al.*<sup>26</sup>

When evaluating derivatives at  $\mathbf{x}_0$ , we need to relate the basis at  $\mathbf{x}_0$  to some basis at  $\mathbf{x} = \mathbf{x}_0 + \Delta\mathbf{x}$ . Suppose the molecular orbitals (MOs) at  $\mathbf{x}_0$  are

$$\phi_m(\mathbf{x}_0) = \sum_{\alpha} \chi_{\alpha}(\mathbf{x}_0) C_{\alpha m}(\mathbf{x}_0), \quad (64)$$

where  $C_{\alpha m}$  are orbital coefficients and  $\chi_{\alpha}$  are atomic orbitals. The unmodified MOs (UMOs) are defined by freezing the orbital coefficients,

$$\phi_m^u(\mathbf{x}) = \sum_{\alpha} C_{\alpha m}(\mathbf{x}_0) \chi_{\alpha}(\mathbf{x}). \quad (65)$$

The UMOs are not orthonormal, however:

$$S_{mn}(\mathbf{x}) = \langle \phi_m^u(\mathbf{x}) | \phi_n^u(\mathbf{x}) \rangle, \quad S_{mn}(\mathbf{x}_0) = \delta_{pq}. \quad (66)$$

Hence, UMOs are related to orthonormalized MOs (or OMOs) through

$$\phi_m(\mathbf{x}) = \sum_n \phi_n^u(\mathbf{x}) T_{nm}(\mathbf{x}), \quad (67)$$

where the connection matrix  $\mathbf{T}(\mathbf{x})$  satisfies  $\mathbf{T}(\mathbf{x}_0) = \mathbf{I}$  and

$$\mathbf{T}(\mathbf{x})^{\dagger} \mathbf{S}(\mathbf{x}) \mathbf{T}(\mathbf{x}) = \mathbf{I}. \quad (68)$$

In the natural connection,  $\mathbf{T}$  is chosen to be

$$\mathbf{T}(\mathbf{x}) = \mathbf{W}(\mathbf{x})^{-1} (\mathbf{W}(\mathbf{x}) \mathbf{S}(\mathbf{x}) \mathbf{W}(\mathbf{x})^{\dagger})^{1/2} = \mathbf{W}(\mathbf{x})^{-1} \Delta(\mathbf{x}), \quad (69)$$



where

$$W_{mn}(\mathbf{x}) = \langle \phi_m^u(\mathbf{x}_0) | \phi_n^u(\mathbf{x}) \rangle. \quad (70)$$

The natural connection minimizes the change in the orthonormalized orbitals at  $\mathbf{x}$  relative to the orbitals at  $\mathbf{x}_0$ .

Let us now relate the orbital space at  $\mathbf{x}$  to the orbital space at  $\mathbf{x}_0$ . In order to do so, we need to consider a *complete* orbital basis (denoted by indices  $pq \dots$ ), which we partition into the OMO basis ( $mn \dots$ ) and the orthogonal complement orbitals, or OCOs ( $uv \dots$ ). For complete bases, we can write

$$\phi_p(\mathbf{x}) = \sum_q \phi_q(\mathbf{x}_0) U_{qp}(\mathbf{x}), \quad U_{qp}(\mathbf{x}) = \langle \phi_q(\mathbf{x}_0) | \phi_p(\mathbf{x}) \rangle. \quad (71)$$

Occupation number states at  $\mathbf{x}$  can thus be expressed as

$$|\Phi(\mathbf{x})\rangle = U(\mathbf{x})|\Phi(\mathbf{x}_0)\rangle, \quad (72)$$

with

$$U(\mathbf{x}) = \exp(-b(\mathbf{x})), \quad b(\mathbf{x}) = \sum_{pq} b_{pq}(\mathbf{x}) a_p^\dagger(\mathbf{x}_0) a_q(\mathbf{x}_0), \quad (73)$$

where  $b(\mathbf{x})$  is the anti-Hermitian operator with  $b_{pq}(\mathbf{x})$  defined such that  $U(\mathbf{x}) = \exp(-b(\mathbf{x}))$ .

The many-body operators can be expanded as

$$a_p^\dagger(\mathbf{x}) = \sum_q a_q^\dagger(\mathbf{x}_0) U_{qp}(\mathbf{x}). \quad (74)$$

To evaluate derivatives with respect to some specific  $x$ , we expand operators about  $\mathbf{x}_0$ ,

$$a_p^\dagger(\mathbf{x}) = a_p^\dagger + a_p^{(1)\dagger}\Delta x + \frac{1}{2}a_p^{(2)\dagger}(\Delta x)^2 + \dots \quad (75)$$

$$b(\mathbf{x}) = b^{(1)}\Delta x + \frac{1}{2}b^{(2)}(\Delta x)^2 + \dots, \quad (76)$$

where

$$b^{(n)} = \sum_{pq} b_{pq}^{(n)} a_p^\dagger a_q. \quad (77)$$

Here we have let  $a_p^\dagger \equiv a_p^\dagger(\mathbf{x}_0)$  and suppressed the  $\mathbf{x}$ -dependence of the derivatives. It will be useful to split operator contributions in the OMO ( $\parallel$ ) and OCO blocks ( $\perp$ ):

$$a_p^\dagger(\mathbf{x}) = \sum_m a_m^\dagger(\mathbf{x}_0) U_{mp}(\mathbf{x}) + \sum_u a_u^\dagger(\mathbf{x}_0) U_{up}(\mathbf{x}) = a_{p\parallel}^\dagger(\mathbf{x}) + a_{p\perp}^\dagger(\mathbf{x}). \quad (78)$$

Let us evaluate

$$f_{IJ} = \langle \Phi_I(\mathbf{x}_0) | \frac{\partial}{\partial x} \Phi_J(\mathbf{x}) \Big|_0. \quad (79)$$

Using Eq. (72), we get

$$f_{IJ} = \langle \Phi_I(\mathbf{x}_0) | \frac{\partial U}{\partial x} \Big|_0 | \Phi_J(\mathbf{x}_0) \rangle = -\langle \Phi_I(\mathbf{x}_0) | b^{(1)} | \Phi_J(\mathbf{x}_0) \rangle. \quad (80)$$

To simplify further, we note that  $U_{mn} = \Delta_{mn}$  is Hermitian in the natural connection. Since the  $U_{uv}$  block can similarly be chosen to be Hermitian, we have<sup>26</sup>

$$b_{mn} = b_{uv} = 0 \implies b_{mn}^{(k)} = b_{uv}^{(k)} = 0 \quad (81)$$

and so

$$f_{IJ} = - \sum_{mn} b_{mn}^{(1)} \langle \Phi_I(\mathbf{x}_0) | a_m^\dagger a_n | \Phi_J(\mathbf{x}_0) \rangle = 0. \quad (82)$$

In general,  $f_{IJ}$  is non-zero with connections other than the natural connection.

Next, we consider the second derivative

$$g_{IJ} = \langle \Phi_I(\mathbf{x}_0) | \frac{\partial^2}{\partial x^2} \Phi_J(\mathbf{x}) \rangle \Big|_0 = \langle \Phi_I(\mathbf{x}_0) | \frac{\partial^2 U}{\partial x^2} \Big|_0 | \Phi_J(\mathbf{x}_0) \rangle, \quad (83)$$

which can be written

$$g_{IJ} = \langle \Phi_I(\mathbf{x}_0) | -b^{(2)} + b^{(1)}b^{(1)} | \Phi_J(\mathbf{x}_0) \rangle = \langle \Phi_I(\mathbf{x}_0) | b^{(1)}b^{(1)} | \Phi_J(\mathbf{x}_0) \rangle. \quad (84)$$

In the final equality, we have used Eq. (81). Now, notice that since

$$b^{(1)} = \sum_{um} b_{um}^{(1)} a_u^\dagger a_m + \sum_{mu} b_{mu}^{(1)} a_m^\dagger a_u = - \sum_m a_{m\perp}^{(1)\dagger} a_m + \sum_m a_m^\dagger a_{m\perp}^{(1)}, \quad (85)$$

the only non-zero  $b^{(1)}b^{(1)}$  contribution is the one that first creates an electron in the complementary space and then destroys it. Thus,

$$\begin{aligned} g_{IJ} &= - \sum_{mn} \langle \Phi_I(\mathbf{x}_0) | a_m^\dagger a_{m\perp}^{(1)} a_{n\perp}^{(1)\dagger} a_n | \Phi_J(\mathbf{x}_0) \rangle \\ &= - \sum_{mn} \langle \Phi_I(\mathbf{x}_0) | a_m^\dagger [a_{m\perp}^{(1)}, a_{n\perp}^{(1)\dagger}]_+ a_n | \Phi_J(\mathbf{x}_0) \rangle. \end{aligned} \quad (86)$$

The commutator can be expressed as

$$[a_{m\perp}^{(1)}, a_{n\perp}^{(1)\dagger}]_+ = \sum_{uv} U_{um}^{(1)*} U_{vn}^{(1)} [a_u, a_v^\dagger]_+ = \sum_u U_{um}^{(1)*} U_{un}^{(1)} = \sum_u \langle \phi_m^{(1)} | \phi_u(\mathbf{x}_0) \rangle \langle \phi_u(\mathbf{x}_0) | \phi_n^{(1)} \rangle. \quad (87)$$

Moreover, since

$$\langle \phi_m(\mathbf{x}_0) | \phi_n^{(1)} \rangle = U_{mn}^{(1)} = -b_{mn}^{(1)} = 0, \quad (88)$$

the inner projection in Eq. (87) is equivalent to the identity and so

$$[a_{m\perp}^{(1)}, a_{n\perp}^{(1)\dagger}]_+ = \langle \phi_m^{(1)} | \phi_n^{(1)} \rangle. \quad (89)$$

Hence, we get the final result

$$g_{IJ} = - \sum_{mn} \langle \Phi_I(\mathbf{x}_0) | a_m^\dagger a_n | \Phi_J(\mathbf{x}_0) \rangle \langle \phi_m^{(1)} | \phi_n^{(1)} \rangle. \quad (90)$$

The formulas for  $f_{IJ}$  and  $g_{IJ}$  are valid for occupation number states but allow for generalization to general wave functions. We will be concerned with evaluating partial derivatives with respect to  $x$  for wave functions of the form

$$|\psi_k(\mathbf{x})\rangle = \sum_I c_{Ik}(\mathbf{x}) |\Phi_I(\mathbf{x})\rangle. \quad (91)$$

Since the  $c_{Ik}$  depend implicitly on  $\mathbf{x}$ , we have  $\partial c_{Ik} / \partial x = 0$ . Thus,

$$\begin{aligned} f_{kl} &= \langle \psi_k(\mathbf{x}_0) | \frac{\partial}{\partial x} | \psi_l(\mathbf{x}) \rangle \Big|_0 \\ &= \sum_{IJ} c_{Ik}^*(\mathbf{x}_0) \langle \Phi_I(\mathbf{x}_0) | \frac{\partial}{\partial x} | \Phi_J(\mathbf{x}) \rangle \Big|_0 c_{Jl}(\mathbf{x}_0) \\ &= \sum_{IJ} c_{Ik}^*(\mathbf{x}_0) f_{IJ} c_{Jl}(\mathbf{x}_0) \\ &= 0 \end{aligned} \quad (92)$$

and

$$\begin{aligned}
g_{kl} &= \langle \psi_k(\mathbf{x}_0) | \frac{\partial^2}{\partial x^2} | \psi_l(\mathbf{x}) \rangle \Big|_0 \\
&= \sum_{IJ} c_{Ik}^*(\mathbf{x}_0) g_{IJc_{Jl}}(\mathbf{x}_0) \\
&= - \sum_{mn} \langle \psi_k(\mathbf{x}_0) | a_m^\dagger a_n | \psi_l(\mathbf{x}_0) \rangle \langle \phi_m^{(1)} | \phi_n^{(1)} \rangle.
\end{aligned} \tag{93}$$

For partial derivatives of the energy, we also have to account for the explicit  $\mathbf{x}$ -dependence of the Hamiltonian. We express the OMO Hamiltonian as

$$H = \sum_{pq} h_{pq}(\mathbf{x}) E_{pq}(\mathbf{x}) + \frac{1}{2} \sum_{pqrs} g_{pqrs}(\mathbf{x}) e_{pqrs}(\mathbf{x}), \tag{94}$$

where both the integrals and the operators depend on  $\mathbf{x}$ . However, the dependence of the operators can be ignored in energy derivatives because matrix elements of occupation number states are constant:

$$\langle \Phi_I(\mathbf{x}) | \Phi_J(\mathbf{x}) \rangle = \langle \Phi_I(\mathbf{x}_0) | U(\mathbf{x})^\dagger U(\mathbf{x}) | \Phi_J(\mathbf{x}_0) \rangle = \langle \Phi_I(\mathbf{x}_0) | \Phi_J(\mathbf{x}_0) \rangle \tag{95}$$

In particular, elements involving  $E_{pq}(\mathbf{x})$  and  $e_{pqrs}(\mathbf{x})$  are linear combinations of such overlaps and therefore give no contributions in energy derivatives.<sup>27</sup> The integrals are related to the UMO basis as

$$h_{pq}(\mathbf{x}) = \sum_{mn} T_{mp}(\mathbf{x})^* h_{mn}^u(\mathbf{x}) T_{nq}(\mathbf{x}) \tag{96}$$

$$g_{pqrs}(\mathbf{x}) = \sum_{mnkl} T_{mp}(\mathbf{x})^* T_{nq}(\mathbf{x})^* g_{mnkl}^u(\mathbf{x}) T_{kr}(\mathbf{x}) T_{ls}(\mathbf{x}). \tag{97}$$

By differentiating  $\mathbf{T}\mathbf{W} = \mathbf{W}^\dagger \mathbf{T}^\dagger$  and  $\mathbf{I} = \mathbf{T}^\dagger \mathbf{S}\mathbf{T}$ , we find that

$$\mathbf{T}^{(1)} = -\mathbf{W}^{(1)}. \tag{98}$$

Consequently, the partial derivative of the Hamiltonian can be written

$$H^{(1)} = H_u^{(1)} - \{W^{(1)}, H\}, \quad (99)$$

where  $H_u^{(1)}$  is the derivative of the UMO Hamiltonian and

$$\{W^{(1)}, H\} = \sum_{pq} j_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} j_{pqrs} e_{pqrs}, \quad (100)$$

where

$$j_{pq} = \sum_m (W_{pm}^{(1)} h_{mq} + h_{pm} W_{mq}^{(1)}) \quad (101)$$

$$j_{pqrs} = \sum_m (W_{pm}^{(1)} g_{mqs} + W_{qm}^{(1)} g_{pms} + g_{pqms} W_{mr}^{(1)} + g_{pqrm} W_{ms}^{(1)}). \quad (102)$$

The  $\mathbf{W}^{(1)}$  matrix, given by

$$W_{pq}^{(1)} = \left. \frac{\partial W_{pq}}{\partial x} \right|_0 = \sum_{\alpha\beta} C_{\alpha p}(\mathbf{x}_0) C_{\beta q}(\mathbf{x}_0) \int \chi_\alpha(\mathbf{x}_0) \left. \frac{\partial \chi_\beta}{\partial x} \right|_0 d\mathbf{r}, \quad (103)$$

is analogous to  $S_{pq}^{(1)} = \partial S_{pq} / \partial x|_0$  in the symmetric connection  $\mathbf{T} = \mathbf{S}^{-1/2}$ .

This concludes our discussion of how the geometry dependence of the many-body operators affects energy derivatives and nonadiabatic coupling elements. We refer the reader to Helgaker and Jørgensen<sup>27</sup> for more details regarding connections and energy derivatives and to Olsen *et al.*<sup>26</sup> for more on the natural connection. In the next section, we derive expressions for the nonadiabatic elements in coupled cluster theory.

## Nonadiabatic coupled cluster couplings in a Lagrangian formalism

To obtain a Lagrangian for the vector coupling, Hohenstein<sup>22</sup> defined a quantity whose first derivatives are identical to components of the vector coupling. Although Hohenstein used it

for configuration interaction theory, the observation generalizes straightforwardly to coupled cluster theory. The quantity is the partially frozen overlap

$$\mathcal{O}_{mn}(\mathbf{x}) = \langle \tilde{\psi}_m(\mathbf{x}_0) | \psi_n(\mathbf{x}) \rangle, \quad (104)$$

in terms of which we have

$$(\mathbf{F}_{mn}^I)_i = \left\langle \tilde{\psi}_m(\mathbf{x}) \left| \frac{d}{dx_i} \psi_n(\mathbf{x}) \right. \right\rangle \Big|_0 = \frac{d}{dx_i} \mathcal{O}_{mn}(\mathbf{x}) \Big|_0, \quad i \in I, \quad (105)$$

and

$$G_{mn}^I = \left\langle \tilde{\psi}_m(\mathbf{x}) \left| \sum_{i \in I} \frac{d^2}{dx_i^2} \psi_n(\mathbf{x}) \right. \right\rangle \Big|_0 = \sum_{i \in I} \frac{d^2}{dx_i^2} \mathcal{O}_{mn}(\mathbf{x}) \Big|_0. \quad (106)$$

For convenience, we write  $i \in I$  to signify that  $x_i$  is one of the three coordinates at nucleus  $I$  ( $x$ ,  $y$ , or  $z$ ). Clearly, the vector and scalar couplings are derivatives of the partially frozen overlap and may therefore be evaluated using a Lagrangian. Note that the overlap  $\mathcal{O}_{mn}(\mathbf{x})$  depends on  $\mathbf{x}_0$ . We suppress this dependency for notational simplicity.

The overlap is expressed in terms of coupled cluster wave functions, which depend on  $\mathbf{x}$  but also on a set of wave function parameters  $\boldsymbol{\lambda}$  (which themselves depend on  $\mathbf{x}$ ). Written out in terms of wave function parameters, the overlap is given by

$$\mathcal{O}_{mn}(\mathbf{x}, \boldsymbol{\lambda}) = \langle \tilde{\psi}_m(\mathbf{x}_0) | \exp(-\kappa) \exp(T) | \mathcal{R}_n \rangle, \quad (107)$$

where

$$\langle \tilde{\psi}_m(\mathbf{x}_0) | = \langle \mathcal{L}_m | \exp(-T) \Big|_0 \quad (108)$$

and

$$\kappa = \sum_{p>q} \kappa_{pq} (E_{pq} - E_{qp}) = \sum_{p>q} \kappa_{pq} E_{pq}^- . \quad (109)$$

The  $\kappa$  operator accounts for orbital rotations, meaning changes in the Hartree-Fock orbitals, where, by definition, we have  $\kappa(\mathbf{x}_0) = 0$ . Following the standard recipe, we add the equations (denoted by  $\mathcal{E}_{mn}$ ) that determine the parameters as constraints with associated Lagrangian multipliers (denoted by  $\boldsymbol{\gamma}$ ),

$$\mathcal{L}_{mn}(\boldsymbol{\lambda}, \mathbf{x}, \boldsymbol{\gamma}) = \mathcal{O}_{mn}(\boldsymbol{\lambda}, \mathbf{x}) + \boldsymbol{\gamma}^T \mathcal{E}_{mn}(\boldsymbol{\lambda}, \mathbf{x}), \quad m \neq n, \quad (110)$$

where  $\boldsymbol{\lambda}$  and  $\boldsymbol{\gamma}$  are determined for every  $\mathbf{x}$  by stationarity:

$$\frac{\partial \mathcal{L}_{mn}}{\partial \gamma_k} = (\mathcal{E}_{mn})_k = 0 \quad (111)$$

$$\frac{\partial \mathcal{L}_{mn}}{\partial \lambda_k} = 0. \quad (112)$$

The derivatives of this Lagrangian are identical to the derivatives of the frozen overlap (since  $\mathcal{E}_{mn} = 0$ ). One advantage of the Lagrangian formalism is that it automatically incorporates the  $2n + 1$  and  $2n + 2$  rules for  $\boldsymbol{\lambda}$  and  $\boldsymbol{\gamma}$ , respectively. In particular,

$$(\mathbf{F}_{mn}^I)_i = \left. \frac{d\mathcal{L}_{mn}}{dx_i} \right|_0 = \left. \frac{\partial \mathcal{L}_{mn}}{\partial x_i} \right|_0, \quad i \in I, \quad (113)$$

where the final equality follows from stationarity, see Eqs. (111) and (112). Denoting partial derivatives with respect to geometrical coordinates as

$$a^{(i)} = \left. \frac{\partial a}{\partial x_i} \right|_0, \quad a^{(i,j)} = \left. \frac{\partial^2 a}{\partial x_i \partial x_j} \right|_0, \quad (114)$$



we can write

$$(\mathbf{F}_{mn}^I)_i = \mathcal{L}_{mn}^{(i)}, \quad i \in I. \quad (115)$$

Furthermore, if we let

$$f_\alpha = \left. \frac{\partial \mathcal{L}_{mn}}{\partial \lambda_\alpha} \right|_{\lambda=\lambda_0}, \quad H_{\alpha\beta} = \left. \frac{\partial^2 \mathcal{L}_{mn}}{\partial \lambda_\alpha \partial \lambda_\beta} \right|_0, \quad (116)$$

then the scalar coupling can be expressed as (see Appendix A)

$$G_{mn}^I = \sum_{i \in I} \left. \frac{d^2 \mathcal{L}_{mn}}{dx_i^2} \right|_0 = \sum_{i \in I} \left( \mathcal{L}_{mn}^{(i,i)} + \sum_{\alpha\beta} \lambda_\alpha^{(i)} H_{\alpha\beta} \lambda_\beta^{(i)} + 2 \sum_{\alpha} \lambda_\alpha^{(i)} f_\alpha^{(i)} \right). \quad (117)$$

Clearly,  $\mathbf{F}_{mn}^I$  and  $G_{mn}^I$  are similar in complexity to the energy gradient and Hessian. However,  $G_{mn}^I$  is somewhat simpler than the energy Hessian because the first derivatives of the parameters ( $\lambda^{(i)}$ ) can be considered one at a time.

To proceed, we must define the Lagrangian  $\mathcal{L}_{mn}$  in detail. The conditions  $\mathcal{E}_{mn}$  include all equations that must be solved to evaluate the overlap  $\mathcal{O}_{mn}$ . These are (a) the Hartree-Fock equations, (b) the amplitude equations, and (c) the eigenvalue equations for the right state amplitudes. Written out in full, we have

$$\begin{aligned} \mathcal{L}_{mn} &= \mathcal{O}_{mn} + \boldsymbol{\gamma}^T \mathcal{E}_{mn} \\ &= \mathcal{O}_{mn} + \bar{\boldsymbol{\kappa}}^T \mathbf{F}_c + \boldsymbol{\zeta}^T \boldsymbol{\Omega} + \boldsymbol{\beta}_n^T (\bar{\mathcal{H}} - E_n) \mathcal{R}_n + \bar{E}_n (1 - \langle \mathcal{L}_n^0 | \mathcal{R}_n \rangle), \end{aligned} \quad (118)$$

where we have introduced multipliers associated with the different sets of equations,  $\bar{\boldsymbol{\kappa}}$ ,  $\boldsymbol{\zeta}$ , as well as  $\boldsymbol{\beta}_n$  and  $\bar{E}_n$ . We have also introduced the Brillouin condition

$$(\mathbf{F}_c)_{pq} = \langle \text{HF} | [E_{pq}^-, H(\kappa)] | \text{HF} \rangle, \quad p > q, \quad (119)$$

where

$$H(\kappa) = \exp(\kappa)H \exp(-\kappa). \quad (120)$$

Furthermore, the similarity transformed Hamiltonian in  $\mathbf{\Omega}$  and  $\bar{\mathcal{H}}$  is given by

$$\bar{H} = \bar{H}(\kappa) = \exp(-T)H(\kappa) \exp(T) \quad (121)$$

and the  $n$ th electronic energy defined as

$$E_n = \langle \mathcal{L}_n^{(0)} | \bar{H} | \mathcal{R}_n \rangle. \quad (122)$$

With  $\mathcal{L}_{mn}$  defined, we can now consider the equations for the zeroth order multipliers. These are determined from the zeroth order terms of the  $\boldsymbol{\lambda}$  stationarity, Eq. (112). To keep our notation simple, we will denote the zeroth order terms as  $\boldsymbol{\gamma}^{(0)} \equiv \boldsymbol{\gamma}$  and  $\boldsymbol{\lambda}^{(0)} \equiv \boldsymbol{\lambda}$ , where it should be understood from context when these are  $\boldsymbol{\gamma}$  and  $\boldsymbol{\lambda}$  evaluated at  $\mathbf{x}_0$ . Differentiation with respect to the state parameters gives

$$\left. \frac{\partial \mathcal{L}_{mn}}{\partial \mathcal{R}_n} \right|_0 = 0 = \boldsymbol{\mathcal{L}}_m^T + \boldsymbol{\beta}_n^T (\bar{\mathcal{H}} - E_n \mathbf{I}) - (\bar{E}_n + E_n \boldsymbol{\beta}_n^T \mathbf{R}_n) \boldsymbol{\mathcal{L}}_n^T = 0. \quad (123)$$

To solve this equation, we note that if we let

$$\bar{E}_n = -E_n \boldsymbol{\beta}_n^T \mathbf{R}_n, \quad (124)$$

the equation for  $\boldsymbol{\beta}_n$  becomes

$$\boldsymbol{\beta}_n^T (\bar{\mathcal{H}} - E_n \mathbf{I}) = -\boldsymbol{\mathcal{L}}_m^T. \quad (125)$$

Thus, we have

$$\boldsymbol{\beta}_n = (E_n - E_m)^{-1} \boldsymbol{\mathcal{L}}_m. \quad (126)$$

Next we consider stationarity with respect to  $\mathbf{t}$ . This can be expressed as

$$\left. \frac{\partial \mathcal{L}_{mn}}{\partial \mathbf{t}} \right|_0 = 0 = {}^t \boldsymbol{\eta} + \boldsymbol{\zeta}^T \mathbf{A}, \quad (127)$$

where

$${}^t \eta_\mu = \langle \bar{\mathcal{L}}_m | \tau_\mu | \bar{\mathcal{R}}_n \rangle + \langle \bar{\beta}_n | [H, \tau_\mu] | \bar{\mathcal{R}}_n \rangle \quad (128)$$

and where we have introduced the notation

$$|\bar{X}\rangle = \exp(T)|X\rangle \quad (129)$$

$$\langle \bar{X}| = \langle X| \exp(-T). \quad (130)$$

Finally, we have stationarity with respect to  $\boldsymbol{\kappa}$ , which can be written

$$\left. \frac{\partial \mathcal{L}_{mn}}{\partial \boldsymbol{\kappa}} \right|_0 = 0 = {}^\kappa \boldsymbol{\eta} + \bar{\boldsymbol{\kappa}}^T \mathbf{A}^{\text{HF}}, \quad (131)$$

where

$${}^\kappa \eta_{rs} = - \langle \bar{\mathcal{L}}_m | E_{rs}^- | \bar{\mathcal{R}}_n \rangle + \langle \bar{\zeta} | [E_{rs}^-, H] | \text{CC} \rangle + \langle \bar{\beta}_n | [E_{rs}^-, H] | \bar{\mathcal{R}}_n \rangle \quad (132)$$

and

$$A_{pq,rs}^{\text{HF}} = \langle \text{HF} | [E_{pq}^-, [E_{rs}^-, H]] | \text{HF} \rangle. \quad (133)$$

With the zeroth order multipliers determined, we can derive the expression for the vector coupling. By partially differentiating  $\mathcal{L}_{mn}$ , we find that

$$\begin{aligned}
(\mathbf{F}_{mn}^I)_i &= \mathcal{L}_{mn}^{(i)} = (E_n - E_m)^{-1} \langle \mathcal{L}_m | \bar{H}^{(i)} | \mathcal{R}_n \rangle \\
&\quad + \langle \zeta | \bar{H}^{(i)} | \text{HF} \rangle + \langle \text{HF} | [\bar{\kappa}, H^{(i)}] | \text{HF} \rangle, \quad i \in I,
\end{aligned}
\tag{134}$$

where

$$\bar{H}^{(i)} = \exp(-T) H^{(i)} \exp(T)
\tag{135}$$

and where quantities at  $\mathbf{x}_0$  are denoted as  $y^{(0)} \equiv y$  (e.g., we denote  $T^{(0)}$  as  $T$ ).

The vector coupling given in Eq. (134) has also been identified by other authors. It was derived by Christiansen,<sup>13</sup> who assumed biorthonormality and used  $Z$ -vector substitution<sup>14</sup> on the expression for the vector coupling. Tajti and Szalay<sup>15</sup> identified the same expression indirectly via  $Z$ -vector substitution on derivatives of Hamiltonian transition elements. However, they also argued<sup>15</sup> that the coupling should not be given by Eq. (134) but rather be averaged and expressed with normalized states. As we have shown, Eq. (134) is a valid choice due to norm invariance and represents the vector coupling in the right nuclear Schrödinger equations. For the left Schrödinger equations, we can make use of the identity

$$\langle \tilde{\psi}_m | \psi_n \rangle = \delta_{mn} \implies \mathbf{F}_{mn}^I = -\tilde{\mathbf{F}}_{mn}^I.
\tag{136}$$

Before moving on to the scalar coupling, we note that although the  $Z$ -vector substitution method is equivalent to the Lagrangian technique, the latter method gives, in our opinion, an especially elegant way of deriving the coupling elements.

For the scalar coupling, we must determine the first derivatives of the parameters. Equations for these are obtained as the first order terms of the multiplier stationarity conditions.

In the case of  $\mathbf{t}$ , we have

$$\left. \frac{d\Omega}{dx_i} \right|_0 = 0 = {}^t \boldsymbol{\xi}^{(i)} + \mathbf{A} \mathbf{t}^{(i)}, \quad (137)$$

where

$${}^t \xi_\mu^{(i)} = \langle \mu | \bar{H}^{(i)} | \text{HF} \rangle + \langle \mu | \exp(-T) [\kappa^{(i)}, H] \exp(T) | \text{HF} \rangle. \quad (138)$$

In the case of  $\boldsymbol{\kappa}$ , we similarly have

$$\left. \frac{d\mathbf{F}_c}{dx_i} \right|_0 = 0 = \boldsymbol{\kappa} \boldsymbol{\xi}^{(i)} + \mathbf{A}^{\text{HF}} \boldsymbol{\kappa}^{(i)}, \quad (139)$$

where

$$\boldsymbol{\kappa}_{rs}^{(i)} = \langle \text{HF} | [E_{rs}^-, H^{(i)}] | \text{HF} \rangle. \quad (140)$$

The binormality condition implies

$$\left. \frac{d}{dx_i} (1 - \langle \mathcal{L}_n^0 | \mathcal{R}_n \rangle) \right|_0 = 0 = -\mathcal{L}_n^T \mathcal{R}_n^{(i)}, \quad (141)$$

while the eigenvalue condition implies

$$\left. \frac{d}{dx_i} \left( (\bar{\mathcal{H}} - E_n) \mathcal{R}_n \right) \right|_0 = 0 = \mathcal{R}_n \boldsymbol{\xi}^{(i)} + (\bar{\mathcal{H}} - E_n) \mathcal{R}_n^{(i)} \quad (142)$$

Here we have defined

$$\mathcal{R}_n \boldsymbol{\xi}^{(i)} = \left( \boldsymbol{\mathcal{Y}}^{(i)} - \mathcal{L}_n^T \boldsymbol{\mathcal{Y}}^{(i)} \mathcal{R}_n \right) \mathcal{R}_n \quad (143)$$

where

$$\mathcal{Y}_{\mu\nu}^{(i)} = \langle \mu | \bar{H}^{(i)} | \nu \rangle + \langle \mu | [\bar{H}, T^{(i)}] | \nu \rangle + \langle \mu | [\kappa^{(i)}, \bar{H}] | \nu \rangle, \quad \mu, \nu \geq 0. \quad (144)$$

With the derivatives of the parameters determined, let us next consider  $f_\alpha^{(i)}$  and  $H_{\alpha\beta}$ , see Eq. (116). Recall that the  $\alpha$  and  $\beta$  indices refer to the parameters  $\lambda_\alpha$  and  $\lambda_\beta$ . The gradient  $\mathbf{f}$  is given by the zeroth order equations for the multipliers, that is, Eqs. (123), (127), and (131), with  $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0$  but allowing for  $\mathbf{x} \neq \mathbf{x}_0$ . Partially differentiating these terms with respect to  $x_i$  gives  $\mathbf{f}^{(i)}$ . The blocks of the  $\sum_\alpha f_\alpha^{(i)} \lambda_\alpha^{(i)}$  contributions to  $G_{mn}^I$  may be written

$$\mathcal{R}_n f_\mu^{(i)} (\mathcal{R}_n^{(i)})_\mu = \left. \frac{\partial^2 \mathcal{L}_{mn}}{\partial x_i \partial \mathcal{R}_\mu^n} \right|_0 (\mathcal{R}_n^{(i)})_\mu = \langle \bar{\beta}_n | H^{(i)} - E_n^{(i)} | \bar{\mathcal{R}}_n \rangle \quad (145)$$

and

$${}^t f_\mu^{(i)} t_\mu^{(i)} = \left. \frac{\partial^2 \mathcal{L}_{mn}}{\partial x_i \partial t_\mu} \right|_0 t_\mu^{(i)} = \langle \bar{\zeta} | [H^{(i)}, T^{(i)}] | \text{CC} \rangle + \langle \bar{\beta}_n | [H^{(i)}, T^{(i)}] | \bar{\mathcal{R}}_n \rangle \quad (146)$$

$$\begin{aligned} \kappa f_{pq}^{(i)} \kappa_{pq}^{(i)} &= \left. \frac{\partial^2 \mathcal{L}_{mn}}{\partial x_i \partial \kappa_{pq}} \right|_0 \kappa_{pq}^{(i)} = \langle \bar{\zeta} | [\kappa^{(i)}, H^{(i)}] | \text{CC} \rangle + \langle \bar{\beta}_n | [\kappa^{(i)}, H^{(i)}] | \bar{\mathcal{R}}_n \rangle \\ &+ \langle \text{HF} | [\bar{\kappa}, [\kappa^{(i)}, H^{(i)}]] | \text{HF} \rangle \end{aligned} \quad (147)$$

where repeated indices implies summation. For contributions to  $G_{mn}^I$  involving the parameter Hessian  $H_{\alpha\beta} = \partial^2 \mathcal{L}_{mn} / \partial \lambda_\alpha \partial \lambda_\beta |_0$ , we have, for terms involving  $\mathbf{t}$  and  $\boldsymbol{\kappa}$ ,

$$\begin{aligned} t_\mu^{(i)} H_{\mu\nu}^{tt} t_\nu^{(i)} &= t_\mu^{(i)} \left. \frac{\partial^2 \mathcal{L}_{mn}}{\partial t_\mu \partial t_\nu} \right|_0 t_\nu^{(i)} \\ &= \langle \bar{\mathcal{L}}_m | (T^{(i)})^2 | \bar{\mathcal{R}}_n \rangle \\ &+ \langle \bar{\zeta} | [[H, T^{(i)}], T^{(i)}] | \text{CC} \rangle \\ &+ \langle \bar{\beta}_n | [[H, T^{(i)}], T^{(i)}] | \bar{\mathcal{R}}_n \rangle \end{aligned} \quad (148)$$

as well as

$$\begin{aligned}
\kappa_{rs}^{(i)} H_{rs\nu}^{\kappa t} t_\nu^{(i)} &= \kappa_{rs}^{(i)} \frac{\partial^2 \mathcal{L}_{mn}}{\partial \kappa_{rs} \partial t_\nu} \Big|_0 t_\nu^{(i)} \\
&= -\langle \bar{\mathcal{L}}_m | \kappa^{(i)} T^{(i)} | \bar{\mathcal{R}}_n \rangle \\
&+ \langle \bar{\zeta} | [[\kappa^{(i)}, H], T^{(i)}] | \text{CC} \rangle \\
&+ \langle \bar{\beta}_n | [[\kappa^{(i)}, H], T^{(i)}] | \bar{\mathcal{R}}_n \rangle
\end{aligned} \tag{149}$$

and

$$\begin{aligned}
\kappa_{pq}^{(i)} H_{pqrs}^{\kappa\kappa} \kappa_{rs}^{(i)} &= \kappa_{pq}^{(i)} \frac{\partial^2 \mathcal{L}_{mn}}{\partial \kappa_{pq} \partial \kappa_{rs}} \Big|_0 \kappa_{rs}^{(i)} \\
&= \langle \bar{\mathcal{L}}_m | (\kappa^{(i)})^2 | \bar{\mathcal{R}}_n \rangle \\
&+ \langle \bar{\zeta} | [\kappa^{(i)}, [\kappa^{(i)}, H]] | \text{CC} \rangle \\
&+ \langle \bar{\beta}_n | [\kappa^{(i)}, [\kappa^{(i)}, H]] | \bar{\mathcal{R}}_n \rangle \\
&+ \langle \text{HF} | [\bar{\kappa}, [\kappa^{(i)}, [\kappa^{(i)}, H]]] | \text{HF} \rangle.
\end{aligned} \tag{150}$$

Next we have terms involving right state and the cluster amplitudes and orbital rotations:

$$\begin{aligned}
t_\mu^{(i)} H_{\mu\nu}^{t\mathcal{R}_n} (\mathcal{R}_n^{(i)})_\nu &= t_\mu^{(i)} \frac{\partial^2 \mathcal{L}_{mn}}{\partial t_\mu \partial (\mathcal{R}_n)_\nu} \Big|_0 (\mathcal{R}_n^{(i)})_\nu \\
&= \langle \bar{\mathcal{L}}_m | T^{(i)} | \bar{\mathcal{R}}_n^{(i)} \rangle \\
&+ \langle \bar{\beta}_n | [H, T^{(i)}] | \bar{\mathcal{R}}_n^{(i)} \rangle \\
&- \langle \beta_n | \mathcal{R}_n^{(i)} \rangle \langle \bar{\mathcal{L}}_n | [H, T^{(i)}] | \bar{\mathcal{R}}_n \rangle
\end{aligned} \tag{151}$$

$$\begin{aligned}
\kappa_{pq}^{(i)} H_{pq\nu}^{\kappa\mathcal{R}_n} (\mathcal{R}_n^{(i)})_\nu &= \kappa_{pq}^{(i)} \frac{\partial^2 \mathcal{L}_{mn}}{\partial \kappa_{pq} \partial (\mathcal{R}_n)_\nu} \Big|_0 (\mathcal{R}_n^{(i)})_\nu \\
&= -\langle \bar{\mathcal{L}}_m | \kappa^{(i)} | \bar{\mathcal{R}}_n^{(i)} \rangle \\
&+ \langle \bar{\beta}_n | [\kappa^{(i)}, H] | \bar{\mathcal{R}}_n^{(i)} \rangle \\
&- \langle \beta_n | \mathcal{R}_n^{(i)} \rangle \langle \bar{\mathcal{L}}_n | [\kappa^{(i)}, H] | \bar{\mathcal{R}}_n \rangle.
\end{aligned} \tag{152}$$

Finally, we have the partial derivative of the Lagrangian, which can be written

$$\begin{aligned}\mathcal{L}_{mn}^{(i,i)} &= - \sum_{pq} \langle \bar{\mathcal{L}}_m | E_{pq} | \bar{\mathcal{R}}_n \rangle \langle \phi_p^{(i)} | \phi_q^{(i)} \rangle + \langle \bar{\zeta} | H^{(i,i)} | \text{CC} \rangle \\ &+ \langle \text{HF} | [\bar{\kappa}, H^{(i,i)}] | \text{HF} \rangle + \langle \bar{\beta}_n | H^{(i,i)} | \bar{\mathcal{R}}_n \rangle,\end{aligned}\tag{153}$$

Written in compact notation, the scalar coupling may be evaluated as

$$\begin{aligned}G_{mn}^I &= \langle \bar{\mathcal{L}}_m | \bar{R}_n^{(i,i)} \rangle + \langle \bar{\zeta} | K^{(i,i)} | \text{CC} \rangle + \langle \bar{\beta}_n | K^{(i,i)} | \bar{\mathcal{R}}_n \rangle \\ &+ \langle \text{HF} | [\bar{\kappa}, J^{(i,i)}] | \text{HF} \rangle + \langle \bar{\beta}_n | L^{(i)} - \langle L^{(i)} \rangle_n | \bar{\mathcal{R}}_n^{(i)} \rangle\end{aligned}\tag{154}$$

where we have let

$$K^{(i,i)} = J^{(i,i)} + 2[J^{(i)}, T^{(i)}] + [[H, T^{(i)}], T^{(i)}]\tag{155}$$

$$J^{(i)} = H^{(i)} + [\kappa^{(i)}, H]\tag{156}$$

$$L^{(i)} = 2(H^{(i)} + [\kappa^{(i)}, H] + [H, T^{(i)}])\tag{157}$$

$$\langle L^{(i)} \rangle_n = \langle \bar{\mathcal{L}}_n | L^{(i)} | \bar{\mathcal{R}}_n \rangle\tag{158}$$

$$J^{(i,i)} = H^{(i,i)} + 2[\kappa^{(i)}, H^{(i)}] + [\kappa^{(i)}, [\kappa^{(i)}, H]]\tag{159}$$

as well as

$$\begin{aligned}\langle \bar{\mathcal{L}}_m | \bar{R}_n^{(i,i)} \rangle &= \langle \bar{\mathcal{L}}_m | (\kappa^{(i)})^2 - 2\kappa^{(i)}T^{(i)} + (T^{(i)})^2 | \bar{\mathcal{R}}_n \rangle - 2\langle \bar{\mathcal{L}}_m | \kappa^{(i)} | \bar{\mathcal{R}}_n^{(i)} \rangle \\ &+ 2\langle \bar{\mathcal{L}}_m | T^{(i)} | \bar{\mathcal{R}}_n^{(i)} \rangle - \sum_{pq} \langle \bar{\mathcal{L}}_m | E_{pq} | \bar{\mathcal{R}}_n \rangle \langle \phi_p^{(i)} | \phi_q^{(i)} \rangle.\end{aligned}\tag{160}$$

Throughout the derivations above, we have considered the off-diagonal coupling elements ( $m \neq n$ ). The diagonal terms can be derived from the Lagrangian

$$\begin{aligned}\mathcal{L}_{nn} &= \mathcal{O}_{nn} + \gamma^T \mathcal{E}_{nn} \\ &= \mathcal{O}_{nn} + \bar{\kappa}^T \mathbf{F}_c + \zeta^T \boldsymbol{\Omega} + \beta_n^T (\bar{\mathcal{H}} - E_n) \mathcal{R}_n + \bar{E}_n (1 - \langle \mathcal{L}_n^0 | \mathcal{R}_n \rangle),\end{aligned}\tag{161}$$



which gives the slightly different  $\mathcal{R}_n$  stationarity condition

$$0 = (1 + \bar{E}_n - E_n \beta_n^T \mathcal{R}_n) \mathcal{L}_n^T + \beta_n^T (\bar{\mathcal{H}} - E_n). \quad (162)$$

Here we again select  $\bar{E}_n$  to make the first term vanish, giving

$$\beta_n^T = \mathcal{L}_n^T. \quad (163)$$

Other than this change, the derivation of the scalar coupling is virtually unchanged. Terms involving differentiation of  $\mathcal{O}_{mn}$  has the left state  $\langle \mathcal{L}_n |$  in the bra instead of  $\langle \mathcal{L}_m |$  (e.g., in the stationarity conditions for the zeroth order multipliers). In particular, the expression in Eq. (154) is valid with  $m = n$ .

Unlike for the vector coupling, there is no convenient relationship between  $G_{mn}^I$  and  $\tilde{G}_{mn}^I$ . To derive the latter quantity, we may consider the Lagrangian

$$\begin{aligned} \mathcal{L}_{mn} &= \mathcal{O}_{mn} + \gamma^T \mathcal{E}_{mn} \\ &= \mathcal{O}_{mn} + \bar{\kappa}^T \mathbf{F}_c + \zeta^T \Omega + \mathcal{L}_m^T (\bar{\mathcal{H}} - E_m) \beta_m + \bar{E}_m (1 - \langle \mathcal{L}_m | \mathcal{R}_m^0 \rangle), \end{aligned} \quad (164)$$

where

$$\mathcal{O}_{mn} = \langle \mathcal{L}_m | \exp(-T) \exp(\kappa) | \psi_n(\mathbf{x}_0) \rangle \quad (165)$$

$$E_m = \mathcal{L}_m^T \bar{\mathcal{H}} \mathcal{R}_m^0. \quad (166)$$

The  $\mathcal{L}_m$  stationarity then gives

$$0 = \mathcal{R}_n + (\bar{\mathcal{H}} - E_m) \beta_m + (\bar{E}_m - E_m \mathcal{L}_m^T \beta_m) \mathcal{R}_m, \quad (167)$$

from which we again have  $\bar{E}_m = E_m \mathcal{L}_m^T \beta_m$  and thus

$$\beta_m = -(E_n - E_m)^{-1} \mathcal{R}_n. \quad (168)$$

The equations for the zeroth order multipliers are derived as before, with the result that the multipliers change their sign, thus giving the result in Eq. (136) for the vector coupling. For the derivative of the parameters, we have the same equations for  $\mathbf{t}^{(i)}$  and  $\boldsymbol{\kappa}^{(i)}$ . For the derivative of  $\mathcal{L}_m$ , we must solve the equation

$$\left. \frac{d}{dx_i} \left( \mathcal{L}_m^T (\bar{\mathcal{H}} - E_m) \right) \right|_0 = 0 = \mathcal{L}_m \boldsymbol{\xi}^{(i)T} + \mathcal{L}_m^{(i)T} (\bar{\mathcal{H}} - E_m), \quad (169)$$

which is analogous to Eq. (142). In contributions involving  $\langle \beta_n | \dots | \mathcal{R}_n \rangle$  in  $G_{mn}^I$ , we obtain similar expressions involving  $\langle \mathcal{L}_m | \dots | \beta_m \rangle$  in the case of  $\tilde{G}_{mn}^I$ . The end-result is

$$\begin{aligned} \tilde{G}_{mn}^I &= \langle \bar{\mathcal{L}}_m^{(i,i)} | \bar{R}_n \rangle + \langle \bar{\zeta} | K^{(i,i)} | \text{CC} \rangle + \langle \bar{\mathcal{L}}_m | K^{(i,i)} | \bar{\beta}_m \rangle \\ &+ \langle \text{HF} | [\bar{\kappa}, J^{(i,i)}] | \text{HF} \rangle + \langle \bar{\mathcal{L}}_m^{(i)} | L^{(i)} - \langle L^{(i)} \rangle_m | \bar{\beta}_m \rangle, \end{aligned} \quad (170)$$

with

$$\begin{aligned} \langle \bar{\mathcal{L}}_m^{(i,i)} | \bar{R}_n \rangle &= \langle \bar{\mathcal{L}}_m | (\kappa^{(i)})^2 - 2T^{(i)} \kappa^{(i)} + (T^{(i)})^2 | \bar{\mathcal{R}}_n \rangle + 2 \langle \bar{\mathcal{L}}_m^{(i)} | \kappa^{(i)} | \bar{\mathcal{R}}_n \rangle \\ &- 2 \langle \bar{\mathcal{L}}_m^{(i)} | T^{(i)} | \bar{\mathcal{R}}_n \rangle - \sum_{pq} \langle \bar{\mathcal{L}}_m | E_{pq} | \bar{\mathcal{R}}_n \rangle \langle \phi_p^{(i)} | \phi_q^{(i)} \rangle. \end{aligned} \quad (171)$$

Finally,  $\tilde{G}_{mn}^I$  is obtained in a similar manner to  $G_{nn}^I$ , see Eq. (161) and the surrounding text.

This concludes our derivation of the coupled cluster scalar coupling. To the best of our knowledge, equations for this coupling have not been presented in the literature before (with  $m \neq n$ ). Diagonal terms were also considered by Gauss *et al.*<sup>18</sup> from a different starting point. The scalar coupling is often omitted in dynamics simulations, but its potential influence on nonadiabatic dynamics has been highlighted in recent years (see Curchod and Martínez<sup>3</sup> and references therein).

## Concluding remarks

The norm of the electronic states changes the value of nonadiabatic coupling elements but does not change the molecular wave function. The biorthonormal formula assumed by Christiansen<sup>13</sup> is therefore a valid choice for coupled cluster dynamics with the appropriate nuclear Schrödinger equations. More generally, we have shown that the wave function is invariant under invertible transformations of the electronic basis. Of course, the biorthonormal couplings are not directly comparable to the coupling elements of an Hermitian method with normalized states, such as CI or full-CI. However, this reflects the basis-dependence of the couplings and not the validity of the biorthonormal formalism.

We therefore derive a set of nuclear Schrödinger equations assuming biorthonormal projection onto the electronic basis. Combined with expressions derived for the vector and scalar couplings, these nuclear Schrödinger equations serve as a starting point for the application of nonadiabatic dynamics methods to coupled cluster theory.

Our derivations have been restricted to standard coupled cluster theory. However, the Lagrangian formalism is easily extended to similarity constrained coupled cluster methods,<sup>11,12</sup> which are suited to describe relaxation through a conical intersection between excited states. The application to ground state intersections is less straightforward, but may be accessible with approaches that use a different reference than the closed-shell Hartree-Fock state.<sup>28</sup>

## Acknowledgement

We thank Todd J. Martínez for enlightening discussions. We also acknowledge funding from the Marie Skłodowska-Curie European Training Network “COSINE - COmputational Spectroscopy In Natural sciences and Engineering”, Grant Agreement No. 765739, and the Research Council of Norway through FRINATEK projects 263110 and 275506.

## Appendix A: Lagrangian derivatives

Here we derive first and second derivatives of the generic Lagrangian

$$\mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}, \boldsymbol{\gamma}) = \mathcal{O}(\boldsymbol{\lambda}, \mathbf{x}) + \boldsymbol{\gamma}^T \mathcal{E}(\boldsymbol{\lambda}, \mathbf{x}) \quad (172)$$

with respect to  $\mathbf{x}$ . The parameters and multipliers both depend on  $\mathbf{x}$  since they are determined, for a given  $\mathbf{x}$ , from the stationarity conditions

$$\frac{\partial \mathcal{L}}{\partial \lambda_k} = 0, \quad \frac{\partial \mathcal{L}}{\partial \gamma_k} = 0. \quad (173)$$

Using Einstein notation, we can write the Taylor expansion of  $\mathcal{L}$  about some  $\mathbf{x}_0$  as

$$\begin{aligned} \mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}, \boldsymbol{\gamma}) = & \mathcal{L}_0 + \frac{\partial \mathcal{L}}{\partial x_k} \Big|_0 \Delta x_k + \frac{1}{2} \Delta x_k \frac{\partial^2 \mathcal{L}}{\partial x_k \partial x_l} \Big|_0 \Delta x_l + \frac{1}{2} \Delta \lambda_k \frac{\partial^2 \mathcal{L}}{\partial \lambda_k \partial \lambda_l} \Big|_0 \Delta \lambda_l \\ & + \Delta \lambda_k \frac{\partial^2 \mathcal{L}}{\partial \lambda_k \partial x_l} \Big|_0 \Delta x_l + \Delta \gamma_k \frac{\partial^2 \mathcal{L}}{\partial \gamma_k \partial x_l} \Big|_0 \Delta x_l + \Delta \lambda_k \frac{\partial^2 \mathcal{L}}{\partial \lambda_k \partial \gamma_l} \Big|_0 \Delta \gamma_l + \dots, \end{aligned} \quad (174)$$

where we have ignored terms of order three or higher in  $\Delta \mathbf{x}$ . These terms do not contribute to the first and second derivatives and are therefore not relevant to the analysis given here.

In the first derivative, only the partial derivative survives,

$$\frac{d\mathcal{L}}{dx_i} \Big|_0 = \frac{\partial \mathcal{L}}{\partial x_i} \Big|_0. \quad (175)$$

This is due to the stationarity conditions, since they ensure that there are no linear terms in  $\Delta \boldsymbol{\lambda}$  and  $\Delta \boldsymbol{\gamma}$  in the Taylor expansion in Eq. (174). In the second derivative, it is convenient to introduce notation for derivatives with respect to particular components of  $\mathbf{x}$ . We let

$$a^{(i)} = \frac{\partial a}{\partial x_i} \Big|_0 \quad (176)$$

$$a^{(i,j)} = \frac{\partial^2 a}{\partial x_i \partial x_j} \Big|_0. \quad (177)$$

Then we can write

$$\left. \frac{d\mathcal{L}}{dx_i} \right|_0 = \mathcal{L}^{(i)}. \quad (178)$$

and

$$\begin{aligned} \left. \frac{d^2\mathcal{L}}{dx_i dx_j} \right|_0 &= \mathcal{L}^{(i,j)} + \lambda_k^{(i)} \left. \frac{\partial^2\mathcal{L}}{\partial\lambda_k \partial\lambda_l} \right|_0 \lambda_l^{(j)} + \lambda_k^{(j)} \left. \frac{\partial^2\mathcal{L}}{\partial\lambda_k \partial x_i} \right|_0 + \lambda_k^{(i)} \left. \frac{\partial^2\mathcal{L}}{\partial\lambda_k \partial x_j} \right|_0 \\ &+ \gamma_k^{(j)} \left. \frac{\partial^2\mathcal{L}}{\partial\gamma_k \partial x_i} \right|_0 + \gamma_k^{(i)} \left. \frac{\partial^2\mathcal{L}}{\partial\gamma_k \partial x_j} \right|_0 + 2\lambda_k^{(i)} \left. \frac{\partial^2\mathcal{L}}{\partial\lambda_k \partial\gamma_l} \right|_0 \gamma_l^{(j)}. \end{aligned} \quad (179)$$

Now,

$$\gamma_l^{(j)} \left( \left. \frac{\partial^2\mathcal{L}}{\partial x_i \partial\gamma_l} \right|_0 + \lambda_k^{(i)} \left. \frac{\partial^2\mathcal{L}}{\partial\lambda_k \partial\gamma_l} \right|_0 \right) = \gamma_l^{(j)} \left. \frac{d\mathcal{E}_l}{dx_i} \right|_0 = 0, \quad (180)$$

by stationarity, so that

$$\left. \frac{d^2\mathcal{L}}{dx_i dx_j} \right|_0 = \mathcal{L}^{(i,j)} + \lambda_k^{(i)} \left. \frac{\partial^2\mathcal{L}}{\partial\lambda_k \partial\lambda_l} \right|_0 \lambda_l^{(j)} + \lambda_k^{(j)} \left. \frac{\partial^2\mathcal{L}}{\partial\lambda_k \partial x_i} \right|_0 + \lambda_k^{(i)} \left. \frac{\partial^2\mathcal{L}}{\partial\lambda_k \partial x_j} \right|_0. \quad (181)$$

To simplify the notation further, we define derivatives with respect to the parameters:

$$f_k = \left. \frac{\partial\mathcal{L}}{\partial\lambda_k} \right|_0, \quad H_{kl} = \left. \frac{\partial^2\mathcal{L}}{\partial\lambda_k \partial\lambda_l} \right|_0. \quad (182)$$

Thus, we get the final expression for the second derivatives:

$$\left. \frac{d^2\mathcal{L}}{dx_i dx_j} \right|_0 = \mathcal{L}^{(i,j)} + \lambda_k^{(i)} H_{kl} \lambda_l^{(j)} + \lambda_k^{(j)} f_k^{(i)} + \lambda_k^{(i)} f_k^{(j)}. \quad (183)$$

## References

- (1) Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. *Annalen der Physik* **1927**, *389*, 457–484.


- (2) Zhu, X.; Yarkony, D. R. Non-adiabaticity: the importance of conical intersections. *Mol. Phys.* **2016**, *114*, 1983–2013.
- (3) Curchod, B. F. E.; Martínez, T. J. Ab Initio Nonadiabatic Quantum Molecular Dynamics. *Chem. Rev.* **2018**, *118*, 3305–3336.
- (4) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291.
- (5) Krylov, A. I. Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker’s Guide to Fock Space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433–462.
- (6) Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular electronic-structure theory*; John Wiley & Sons, 2014.
- (7) Loos, P.-F.; Scemama, A.; Jacquemin, D. The Quest for Highly Accurate Excitation Energies: A Computational Perspective. *J. Phys. Chem. Lett.* **2020**, *11*, 2374–2383.
- (8) Hättig, C. Structure optimizations for excited states with correlated second-order methods: CC2 and ADC (2). *Advances in quantum chemistry* **2005**, *50*, 37–60.
- (9) Köhn, A.; Tajti, A. Can coupled-cluster theory treat conical intersections? *J. Chem. Phys.* **2007**, *127*, 044105.
- (10) Kjøenstad, E. F.; Myhre, R. H.; Martínez, T. J.; Koch, H. Crossing conditions in coupled cluster theory. *J. Chem. Phys.* **2017**, *147*, 164105.
- (11) Kjøenstad, E. F.; Koch, H. Resolving the Notorious Case of Conical Intersections for Coupled Cluster Dynamics. *J. Phys. Chem. Lett.* **2017**, *8*, 4801–4807.
- (12) Kjøenstad, E. F.; Koch, H. An Orbital Invariant Similarity Constrained Coupled Cluster Model. *J. Chem. Theory Comput.* **2019**, *15*, 5386–5397.

- (13) Christiansen, O. First-order nonadiabatic coupling matrix elements using coupled cluster methods. I. Theory. *J. Chem. Phys.* **1999**, *110*, 711–723.
- (14) Handy, N. C.; Schaefer III, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031–5033.
- (15) Tajti, A.; Szalay, P. G. Analytic evaluation of the nonadiabatic coupling vector between excited states using equation-of-motion coupled-cluster theory. *J. Chem. Phys.* **2009**, *131*, 124104.
- (16) Ichino, T.; Gauss, J.; Stanton, J. F. Quasidiabatic states described by coupled-cluster theory. *J. Chem. Phys.* **2009**, *130*, 174105.
- (17) Purvis, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.
- (18) Gauss, J.; Tajti, A.; Kállay, M.; Stanton, J. F.; Szalay, P. G. Analytic calculation of the diagonal Born-Oppenheimer correction within configuration-interaction and coupled-cluster theory. *J. Chem. Phys.* **2006**, *125*, 144111.
- (19) Faraji, S.; Matsika, S.; Krylov, A. I. Calculations of non-adiabatic couplings within equation-of-motion coupled-cluster framework: Theory, implementation, and validation against multi-reference methods. *J. Chem. Phys.* **2018**, *148*, 044103.
- (20) Born, M.; Huang, K. *Dynamical theory of crystal lattices*; Clarendon press, 1954.
- (21) Shamasundar, K. R. Diagonal Born–Oppenheimer correction for coupled-cluster wavefunctions. *Mol. Phys.* **2018**, *116*, 1483–1495.
- (22) Hohenstein, E. G. Analytic formulation of derivative coupling vectors for complete active space configuration interaction wavefunctions with floating occupation molecular orbitals. *J. Chem. Phys.* **2016**, *145*, 174110.

- (23) Stanton, J. F.; Bartlett, R. J. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* **1993**, *98*, 7029–7039.
- (24) Koch, H.; Jørgensen, P. Coupled cluster response functions. *J. Chem. Phys.* **1990**, *93*, 3333–3344.
- (25) Pacher, T.; Mead, C. A.; Cederbaum, L. S.; Köppel, H. Gauge theory and quasidiabatic states in Mol. Phys. *J. Chem. Phys.* **1989**, *91*, 7057–7062.
- (26) Olsen, J.; Bak, K. L.; Ruud, K.; Helgaker, T.; Jørgensen, P. Orbital connections for perturbation-dependent basis sets. *Theoretica chimica acta* **1995**, *90*, 421–439.
- (27) Helgaker, T.; Jørgensen, P. In *Methods in Computational Molecular Physics*; Wilson, S., Diercksen, G. H. F., Eds.; Springer US: Boston, MA, 1992; pp 353–421.
- (28) Lefrancois, D.; Tuna, D.; Martínez, T. J.; Dreuw, A. The Spin-Flip Variant of the Algebraic-Diagrammatic Construction Yields the Correct Topology of S1/S0 Conical Intersections. *J. Chem. Theory Comput.* **2017**, *13*, 4436–4441.



$$\Psi(\mathbf{r}, \mathbf{R}, t) = \sum_n \chi_n(\mathbf{R}, t) \psi_n(\mathbf{r}; \mathbf{R})$$

$$\tilde{\Psi}(\mathbf{r}, \mathbf{R}, t) = \sum_n \tilde{\chi}_n(\mathbf{R}, t) \tilde{\psi}_n(\mathbf{r}; \mathbf{R})$$


$$(i \frac{d}{dt} - \mathbf{E}) \chi = \sum_I \frac{1}{2M_I} (\mathbf{I} \nabla_I^2 + \mathbf{G}_I + 2\mathbf{F}_I \cdot \nabla_I) \chi$$

$$(-i \frac{d}{dt} - \mathbf{E}) \tilde{\chi} = \sum_I \frac{1}{2M_I} (\mathbf{I} \nabla_I^2 + \tilde{\mathbf{G}}_I + 2\tilde{\mathbf{F}}_I \cdot \nabla_I) \tilde{\chi}$$