A fast method for electronic couplings in embedded multichromophoric systems

OPEN ACCESS
IOP Publishing
Journal of Physics: Condensed Matter
J. Phys.: Condens. Matter **34** (2022) 304004 (11pp)
https://doi.org/10.1088/1361-648X/ac6f3c

A fast method for electronic couplings in embedded multichromophoric systems

Edoardo Cignoni*, Lorenzo Cupellini* and Benedetta Mennucci*

Dipartimento di Chimica e Chimica Industriale, University of Pisa, via G. Moruzzi 13, 56124, Pisa, Italy

E-mail: edoardo.cignoni@phd.unipi.it, lorenzo.cupellini@unipi.it and benedetta.mennucci@unipi.it

Received 17 February 2022, revised 5 May 2022
Accepted for publication 12 May 2022
Published 26 May 2022

Abstract

Electronic couplings are key to understanding exciton delocalization and transport in natural and artificial light harvesting processes. We develop a method to compute couplings in multichromophoric aggregates embedded in complex environments without running expensive quantum chemical calculations. We use a transition charge approximation to represent the quantum mechanical transition densities of the chromophores and an atomistic and polarizable classical model to describe the environment atoms. We extend our framework to estimate transition charges directly from the chromophore geometry, i.e., bypassing completely the quantum mechanical calculations using a regression approach. The method allows to rapidly compute accurate couplings for a large number of geometries along molecular dynamics trajectories.

Keywords: polarizable MM embedding, light harvesting, electronic couplings, excitons, excitation energy transfer

Supplementary material for this article is available online

(Some figures may appear in colour only in the online journal)

1. Introduction

Excitation energy transfer (EET) and delocalization are fundamental mechanisms for natural and artificial light harvesting (LH) processes. In natural LH, both EET efficiency and exciton delocalization concur to shaping the energy transport from antenna complexes to reaction centers [1, 2]. In artificial LH analogs, exciton transport is an important parameter to optimize [3, 4].

The key ingredients for the prediction of energy transfer dynamics are the electronic couplings between the chromophores forming the multichromophoric aggregate. The coupling strength in fact determines the type of EET regime, from incoherent hopping of a localized excitation all the way to wavelike coherent motion of delocalized excitons [5–7]. Moreover, a quantitative modeling of couplings is also required for an accurate simulation of linear and nonlinear electronic spectroscopies [8, 9].

Quantum mechanical (QM) methods now represent key tools for the calculation of electronic couplings [10–12]. In the most common case where bright excitations are involved and the Coulomb interaction is the dominant component of couplings, a QM calculation of the transition density of each chromophore of the aggregate is needed. Moreover, the multichromophoric aggregates are usually embedded in a complex environment, such a protein scaffold or an organic film, which shapes both their properties and interactions. Including the environment effects in the QM calculations is thus mandatory. A very effective way to achieve such a goal is to interface QM methods with classical descriptions of the surrounding environment in the so-called QM/classical multiscale methods.

While QM/classical methods can represent a very accurate strategy for calculating couplings, they are also compu-

* Authors to whom any correspondence should be addressed.

tationally expensive. This is particularly true for antenna complexes whose nuclear dynamics causes fluctuations of the couplings. As a consequence, the entire thermal distribution of couplings has to be characterized. In practice, this means that the QM/classical coupling calculations have to be repeated for many different configurations of the whole system usually generated through a molecular dynamics (MD) simulation. In these cases, inexpensive methods are required to properly sample configurations and obtain exhaustive distributions of couplings.

A possible strategy to reduce the computational cost is to represent the QM transition density as a small set of point charges placed on the chromophore's atoms [13, 14]. These methods based on transition charges have also been interfaced with continuum models of the solvent for describing the dielectric screening of couplings [15, 16]. When applied to MD simulations, in order to avoid computing the QM transition density for many frames, these charges are determined once and then used throughout the simulation [17–24]. In this way, however, they do not account for the effect of the environment or the dynamics. Clearly, a framework that incorporates the dynamics into an inexpensive method for computing the electronic couplings is highly desirable.

Here we propose a fast but accurate method to compute electronic couplings along a MD simulation. First, we extend the transition charges method to be used within an atomistic polarizable description of the environment, which is desirable when dealing with pigments embedded in a protein matrix. In this work, the atoms of the environment are modeled as a set of atomic charges and isotropic polarizabilities. Within this formulation the QM density polarizes the environment generating induced point dipoles. The latter, together with the point charges, induce an electric field which polarizes back the QM density. This polarizable model is commonly known as MMPol and has been successfully used in combination with QM descriptions to model EET processes and other light-induced processes in complex systems [25, 26]. Here, we show that the MMPol model integrated with the transition charge approximation of the transition density well reproduces the exact QM/MMpol coupling. We further employ a regression approach to predict the transition charges directly from the chromophore geometry, bypassing the cost associated with computing the transition densities, and obtaining the vacuum couplings essentially for free. Finally, we leverage the peculiarities of the chromophore transition densities to predict electronic couplings in environment without the need of QM calculations.

We demonstrate and apply the method to the chlorophyll (Chl) molecules in the light-harvesting complex II (LHCII) pigment-protein complex of photosystem II, the major antenna in plants (figure 1(a)). Our regression approach proves to be an inexpensive and accurate way to estimate electronic couplings in the disordered environment of a LH complex.

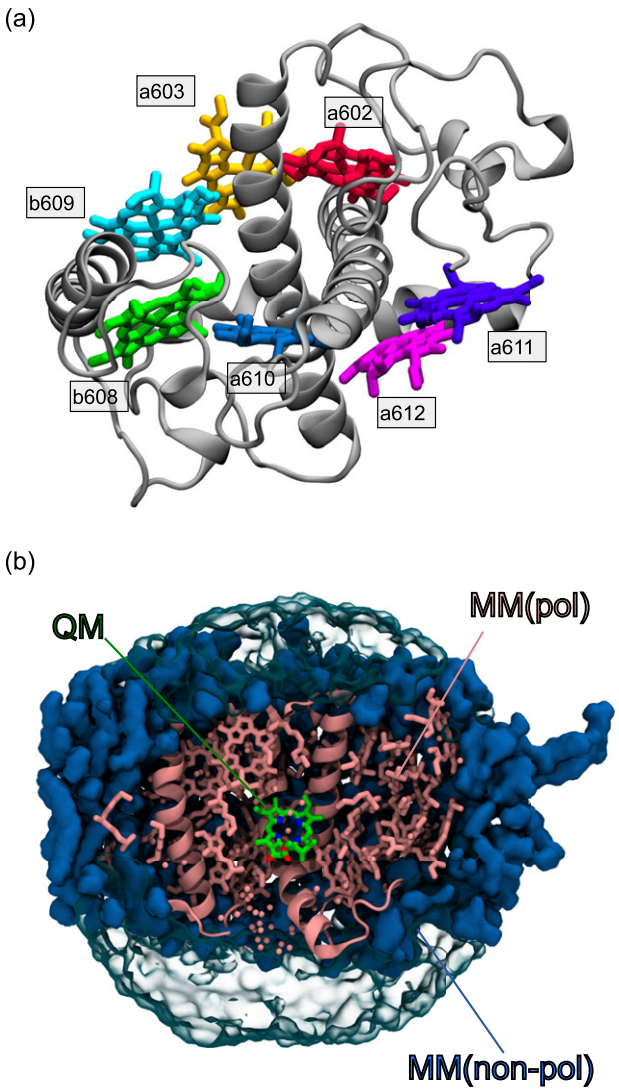

Figure 1. (a) LHCII monomer and the selection of chlorophylls considered in this work. (b) Example of the partition employed for QM/MMPol calculations. The QM subsystem is shown in green. The polarizable MM part is shown in pink. The non-polarizable MM part is shown in blue.

2. Methods and computational details

2.1. The electronic coupling

In the absence of a surrounding environment, the electronic coupling between two interacting pigments can be decomposed in the sum of a long-range Coulomb contribution V_{Coul} and a short-range contribution V_{short} [10, 12, 27]. When bright transitions are considered, the Coulomb term is the dominant contribution. The short-range contribution in fact becomes important only when there is significant overlap between the molecular orbitals of the pigments, i.e., when they are in close proximity. Throughout this paper, the short-range term

is neglected, and the total coupling is approximated with the Coulomb contribution, $V_{\mathrm{Tot}} = V_{\mathrm{Coul}}$.

For two pigments (here indicated as i and j) in vacuum, V_{Coul} is a Coulomb interaction between transition densities ρ_m^{tr}, $m = i, j$, associated with the electronic transitions of interest:

$$V^{\rho}_{\mathrm{Coul},ij} = \int\int \frac{\rho_i^{\mathrm{tr}}(\mathbf{r})\,\rho_j^{\mathrm{tr}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \tag{1}$$

where the superscript ρ indicates that this term is computed from the full transition densities.

When the two pigments are embedded in an environment, the electronic coupling is affected in two distinct ways [25, 28–30]. The first effect is an indirect contribution of the environment, which influences the transition densities ρ_m^{tr} of the pigments, and therefore the strength of their Coulomb coupling. Usually, this indirect contribution results in an increase of $V_{\mathrm{Coul},ij}$. The second (direct) effect arises from the environment being able to polarize in response to the electric fields generated by the transition densities [25, 31].

The direct contribution of the environment can only be defined for a polarizable environment. It is expressed as the interaction between the transition density of one pigment and the environment polarization induced by the transition density of the other pigment.

In the polarizable MMPol model used here, this contribution reads [25]:

$$V^{\rho}_{\mathrm{MMPol},ij} = -\sum_k \int \mathrm{d}\mathbf{r}\,\rho_i^{\mathrm{tr}}(\mathbf{r})\frac{(\mathbf{r}-\mathbf{r}_k)}{|\mathbf{r}-\mathbf{r}'|^3}\boldsymbol{\mu}_k^{\mathrm{MMPol}}\left(\rho_j^{\mathrm{tr}}\right) \tag{2}$$

where k runs over the polarizable atoms in the environment. In (2), the dipoles $\boldsymbol{\mu}_k^{\mathrm{MMPol}}\left(\rho_j^{\mathrm{tr}}\right)$ are induced in response to the charge-density ρ_j^{tr}. We note that (2) is symmetrical with respect to the exchange of pigments. The direct contribution of the environment generally has an opposite sign to V_{Coul}, resulting in a screening effect. The total coupling in the presence of the environment is finally expressed as

$$V^{\rho}_{\mathrm{Tot}} = V^{\rho}_{\mathrm{Coul}} + V^{\rho}_{\mathrm{MMPol}}. \tag{3}$$

It is possible to separately quantify both the indirect and direct effect of the environment on the coupling. By denoting the Coulomb coupling in vacuum and in environment as $V^{\mathrm{vac}}_{\mathrm{Coul}}$ and $V^{\mathrm{env}}_{\mathrm{Coul}}$, respectively, the indirect effect is quantified by the following ratio:

$$f^{\mathrm{ind}} = \frac{V^{\mathrm{env}}_{\mathrm{Coul}}}{V^{\mathrm{vac}}_{\mathrm{Coul}}}. \tag{4}$$

This factor is sensitive to the specific environment of each Chl. Hence, it is different, in principle, for each geometry of the full complex, even when considering a single Chl pair. The direct contribution of the environment can be rationalized, following previous works [25, 29], using the ratio between the full coupling and the Coulomb coupling in environment:

$$s = \frac{V^{\mathrm{env}}_{\mathrm{Coul}} + V^{\mathrm{env}}_{\mathrm{MMPol}}}{V^{\mathrm{env}}_{\mathrm{Coul}}}. \tag{5}$$

An efficient and accurate approximation of the Coulomb coupling can be obtained by projecting the ρ_m^{tr} over a set of (transition) charges placed over the atoms of the pigments. In the present work, the transition charges are obtained by fitting the electrostatic potential generated by ρ_m^{tr}. These charges are known as transition charges from electrostatic potentials (TrEsp) [13]. Within this approximation V_{Coul} is expressed as the Coulomb interaction between point charge distributions:

$$V^{\mathrm{TrEsp}}_{\mathrm{Coul},ij} = \sum_{I\in i}\sum_{J\in j}\frac{q_I^{\mathrm{tr}}q_J^{\mathrm{tr}}}{|\mathbf{r}_I-\mathbf{r}_J|} \tag{6}$$

where q_M^{tr} denotes the TrEsp charge of the Mth atom of pigment m, and where $\mathbf{r}_M$ is the position vector of atom M.

In this contribution, we generalize the TrEsp approximation to be used with a MMPol description of the environment, which we call TrEsp–MMPol. Under the TrEsp–MMPol approximation, the transition density of each pigment is described through TrEsp transition charges, whereas the environment is described as a set of polarizable atoms. By substituting the TrEsp representation of ρ_m^{tr} in (2), we obtain:

$$V^{\mathrm{TrEsp}}_{\mathrm{MMPol},ij} = -\sum_k\sum_{I\in i} q_I^{\mathrm{tr}}\frac{(\mathbf{r}_I-\mathbf{r}_k)}{|\mathbf{r}_I-\mathbf{r}_k|^3}\boldsymbol{\mu}_k^{\mathrm{MMPol}}\left(\left\{q^{\mathrm{tr}}\right\}_j\right) \tag{7}$$

where $\left\{q^{\mathrm{tr}}\right\}_m$ represents the set of TrEsp charges placed on the atoms of pigment m, $m = i, j$. Namely, the direct contribution of the environment is expressed as an interaction between TrEsp charges on one pigment and the environment dipoles induced by the electric field generated by the distribution of TrEsp charges on the other pigment. The total TrEsp coupling in environment is thus $V^{\mathrm{TrEsp}}_{\mathrm{Tot}} = V^{\mathrm{TrEsp}}_{\mathrm{Coul}} + V^{\mathrm{TrEsp}}_{\mathrm{MMPol}}$ in analogy with (3).

2.2. Regression models

Our regression task is to predict the TrEsp charges $\mathbf{q}^{\mathrm{tr}} \in \mathbb{R}^{N\times K}$ starting from a set of molecular encodings $\boldsymbol{\chi} \in \mathbb{R}^{N\times D}$, where N is the number of samples, K is the number of TrEsp charges, and D is the dimensionality of the molecular descriptor. In a notation closer to what is usually presented in the statistical literature, the set of molecular encodings $\boldsymbol{\chi}$ constitutes the design matrix $\mathbf{X}$, and the set of TrEsp charges $\mathbf{q}^{\mathrm{tr}}$ the target $\mathbf{Y}$. Here $\mathbf{q}^{\mathrm{tr}}$ denotes the set of TrEsp charges obtained by fitting the electrostatic potential that results from a time-dependent density functional theory (TD-DFT) calculation, as explained in the previous section. Details on the choice of the molecular encoding are presented later in this section.

As a regression model we have employed simple Tikhonov-regularized linear regression, also known as ridge regression. Within this framework, the TrEsp charges are expressed as a linear combination of the inputs $\hat{\boldsymbol{q}}^{\mathrm{tr}} = \boldsymbol{\omega}_0 + \boldsymbol{\chi}\boldsymbol{\omega}$. Here $\boldsymbol{\omega}_0$ is an intercept term estimated as the average over the N frames of the target charges, $\boldsymbol{\omega}_0 = \langle\mathbf{q}^{\mathrm{tr}}\rangle_{\mathrm{frames}}$. We use the 'hat' notation here to distinguish the target TrEsp charges $\mathbf{q}^{\mathrm{tr}}$ from the charges $\hat{\mathbf{q}}^{\mathrm{tr}}$ predicted by the model. After centering the input matrix $\boldsymbol{\chi}$ so that its columns have mean zero, the coefficients $\boldsymbol{\omega} \in \mathbb{R}^{D\times K}$ are obtained as the minimizer of the following loss function:

$$\mathcal{L}(\boldsymbol{\omega}) = \left(\mathbf{q}^{\mathrm{tr}} - \boldsymbol{\omega}_0 - \boldsymbol{\chi}\boldsymbol{\omega}\right)^{\mathrm{T}}\left(\mathbf{q}^{\mathrm{tr}} - \boldsymbol{\omega}_0 - \boldsymbol{\chi}\boldsymbol{\omega}\right) + \lambda\|\boldsymbol{\omega}\|^2 \tag{8}$$

where λ is a Lagrange multiplier that serves to regularize the least squares estimate, helping to prevent overfitting. We have determined the value of λ with leave-one-out cross-validation, which has an analytic expression for ridge regression.

We have also tested the performance of a more powerful model, namely kernel ridge regression (KRR). KRR and its Bayesian generalization Gaussian process regression (GPR) [32] are widely used regression frameworks in quantum machine learning [33]. Within KRR the TrEsp charges are expressed as a linear combination of kernel evaluations $\hat{\boldsymbol{q}}^{\mathrm{tr}} = \mathbf{Ka}$, where $\mathbf{K} \in \mathbb{R}^{N\times N}$ is the Gram matrix (also called 'kernel') of some feature vector $\phi(\boldsymbol{\chi})$ of the inputs that introduces nonlinearity in the regression, $\mathbf{K}_{ij} = \kappa\left(\boldsymbol{\chi}_i, \boldsymbol{\chi}_j\right) = \phi\left(\boldsymbol{\chi}_i\right)^{\mathrm{T}} \phi\left(\boldsymbol{\chi}_j\right)$. The matrix $\mathbf{a} \in \mathbb{R}^{N\times K}$ is the minimizer of the following 'dual' loss function:

$$\mathcal{L}(\mathbf{a}) = \mathbf{a}^{\mathrm{T}}\mathbf{KKa} - 2\mathbf{a}^{\mathrm{T}}\mathbf{Ky} + \lambda\mathbf{a}^{\mathrm{T}}\mathbf{Ka} \tag{9}$$

where λ is a Lagrange multiplier with the same regularizing function as the one in (8). The power of KRR arises from the fact that one does not have to know the mapping $\phi(\boldsymbol{\chi})$ explicitly, as long as the kernel is a positive definite function of its inputs. In this work we have employed two commonly used kernel functions, namely the infinitely differentiable isotropic radial basis function kernel:

$$\kappa\left(\boldsymbol{\chi}_i, \boldsymbol{\chi}_j; l\right) = \exp\left(-\frac{d^2}{2l}\right) \tag{10}$$

and the Matern kernel:

$$\kappa\left(\boldsymbol{\chi}_i, \boldsymbol{\chi}_j; l, \nu\right) = \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{\sqrt{2\nu}d}{l}\right)^{\nu} K_{\nu}\left(\frac{\sqrt{2\nu}d}{l}\right) \tag{11}$$

where $d = \|\boldsymbol{\chi}_i - \boldsymbol{\chi}_j\|$, l and ν are parameters of the kernel, Γ is the Gamma function, and K_ν is the modified Bessel function of second kind. We have employed $\nu = 2.5$ for the Matern kernel. The length scale parameter l and the regularization parameter λ were determined with five-fold cross-validation using the coefficient of determination of the TrEsp charges as a test metric, where for each fold the parameters were optimized with 50 rounds of Bayesian optimization [34–36] using Gaussian process as inference model and the expected improvement [35] as acquisition function.

Within our regression framework, the molecular geometry is encoded through the well-known Coulomb matrix (CM) descriptor [37, 38]:

$$\chi_{\mathrm{CM},ij} = \frac{Z_i Z_j}{r_{ij}} \quad \forall\, i \neq j \tag{12}$$

where Z_i denotes the atomic number of the ith atom, and r_{ij} is the distance between atoms i and j. We exclude here the diagonal part of the CM, defined as $0.5Z_i^{2.4}$ [37], as it is an uninformative constant in the present case. Therefore, only the upper diagonal part of the full CM, diagonal excluded, is included in our descriptor.

The CM can be categorized as a global descriptor, and as such is expected to work well when excitation quantities are to be computed [39–41]. Being defined in terms of (inverse) pairwise distances, the CM descriptor is naturally invariant to global translations and rotations of the molecule. However, it is not invariant to permutations of identical atoms [42–44], each atom being associated to a single row/column of the matrix. We have then modified the CM through permutations of rows and columns corresponding to groups of identical atoms according to their L2 norm, thus incorporating the permutation invariance at the level of the geometry descriptor. We further note that each permutation in the CM descriptor is applied also to the target TrEsp charges for the purpose of fitting the regression model, and an inverse-permutation is applied after prediction to recover the correct order.

2.3. Computational details of the QM calculations

We have focused the QM analysis on a restricted set of chlorophylls of the LHCII monomer, from strongly down to weakly interacting pairs, and involving both chlorophylls *a* and *b*. The set of chlorophylls considered comprises *a*603, *a*603, *b*608, *b*609, *a*610, *a*611, and *a*612, and is shown in figure 1(a). Chlorophyll geometries were extracted from a classical MD simulation of the trimeric form of LHCII in a 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) phospholipid membrane previously run by some of the authors [45, 46], by selecting 40 frames separated by 20 ns. The electronic couplings between pairs of chlorophylls separated by more than 22 were not computed as they are expected to be negligible. The same frames and geometries were employed to run analogous calculations where the environment was included with the MMPol model. In the QM/MMPol description, all the atoms in the simulation box (except the QM part) were included in the MM region, namely the carotenoids, the protein, membrane lipids and water molecules. The MM region extended up to 30 from the QM region and a radius of 15 was used for the polarization cutoff (see figure 1(b)). The MMPol atoms were described using charge and polarizability parameters derived by Wang *et al* [47]. The Coulomb and direct environment contributions were computed both using (1) and (2) [25, 29] and under the TrEsp–MMPol approximation (6) and (7). In all calculations we have employed TD-DFT with the M062X/6-31G(d) level of QM theory as implemented in Gaussian [48] and we have focused on the lowest, bright, excitation of each chlorophyll (namely the Q_y excitation). This functional has been employed to model the excitations in several LHCs [49, 50], and the exciton couplings between Chls have been shown to be weakly dependent on the choice of the DFT functional [49]. Following this procedure, a total number of 560 unique couplings, both in vacuum and in environment, were computed and analyzed.

2.4. Training dataset and data preprocessing

The regression models (one for Chl *a* and one for Chl *b*) have been trained on an expanded training dataset that comprises every chlorophyll in each of the three monomers of LHCII, for a total of 5440 training samples for Chl *a* and 4080 training samples for Chl *b*. Each frame used to extract the chlorophyll geometries in the training dataset is separated by at least 10 ns from the others. The calculations detailed

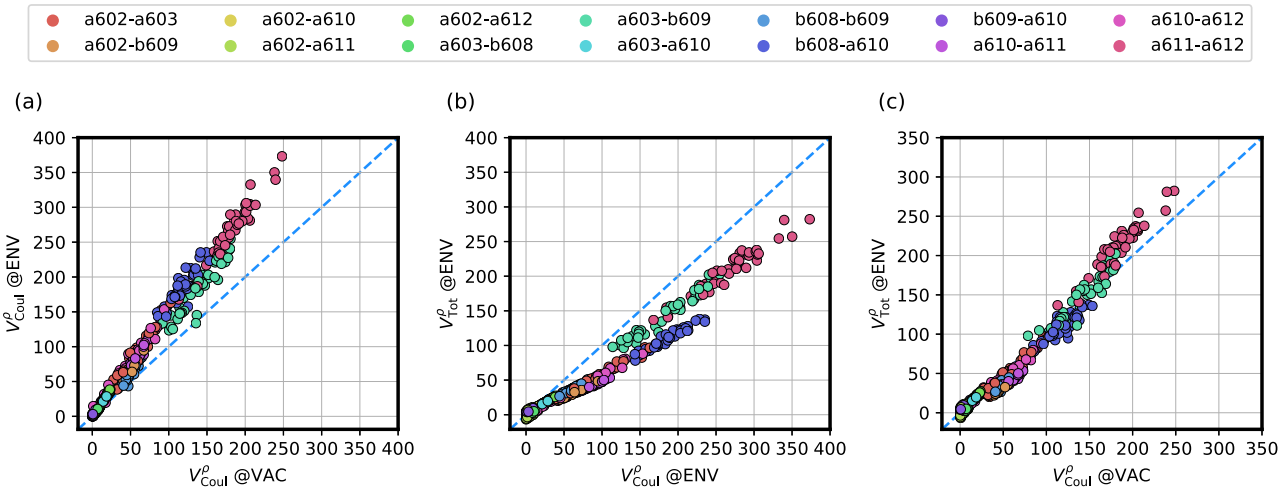

Figure 2. Comparison of various parts of the electronic coupling V in vacuum (@VAC) and in environment (@ENV). (a) Coulomb coupling V_{Coul} in vacuum versus in environment. (b) Coulomb coupling V_{Coul} in environment versus the total coupling V_{Tot} in environment. (c) Coulomb coupling V_{Coul} in vacuum versus the total coupling V_{Tot} in environment. Several Chl pairs have been considered (top legend). All values are in cm^{-1}.

in section 2.3 were not included in the training dataset, and were employed as test dataset for the regression models. The TrEsp charges have been computed in vacuum with the same level of theory (TD-DFT with M062X/6-31G(d)) employed for the test calculations. Regression in quantum machine-learning often has to deal with the arbitrary phase of the wave-function arising from a QM calculation [51–53]. Here we have corrected the sign of the TrEsp charges so that the transition dipole $\boldsymbol{\mu}_m^{\mathrm{tr,TrEsp}} = \sum_{M\in m} q_M^{\mathrm{tr}} \mathbf{r}_M$ makes an acute angle with the NB–ND vector of the chlorophyll ring.

3. Results and discussion

This section is divided as follows: we first analyze in section 3.1 the direct and indirect effects of the composite environment (protein, cofactors, membrane and solvent) on the Chl–Chl couplings in LHCII. In section 3.2 we validate the TrEsp approximation for couplings in vacuum and in the environment. Finally, in sections 3.3 and 3.4 we combine a regression approach and the coupling rescaling to achieve a fast calculation of couplings in the environment.

3.1. Analysis of the electronic couplings in vacuum and in environment

In this section we investigate the relationship between the coupling contributions in vacuum and in environment, to understand their relative importance. The relationships between the Coulomb coupling V^{ρ}_{Coul} in vacuum and in environment and the total coupling V^{ρ}_{Tot} containing the direct and indirect effects of the polarizable environment are shown in figure 2. For an easier interpretation of the results, we set the sign of V^{ρ}_{Coul} as always positive, and change the sign of $V^{\rho}_{\mathrm{MMPol}}$ accordingly.

Figure 2(a) shows the comparison of the Coulomb coupling V^{ρ}_{Coul} in vacuum and in environment. The protein environment enhances the Coulomb coupling, so that V^{ρ}_{Coul} in environment is almost invariably greater than the coupling in vacuum. The increase in Coulomb coupling due to the indirect effect of the environment on the QM transition density can be quantified by the coupling ratio f^{ind} defined in (4). For the Chls pairs considered in this work, f^{ind} varies between 1.38 and 1.73 (Chl *a*–*a* pair), 1.21 and 1.67 (Chl *a*–*b* pair), and 1.11 and 1.38 (Chl *b*–*b* pair). The range of f^{ind} reported shows that the environment influences the transition density of Chls *a* more than it does for Chls *b*. This factor also depends on the orientation and distance of the Chl pair (table 1).

Table 1. f^{ind} factors (4) for the selection of Chl pairs in LHCII. The fifth, 50th, and 95th quantiles of the f^{ind} distribution of each Chl–Chl pair are also reported. The factor is computed for points with vacuum $V^{\rho}_{\mathrm{Coul}} > 10\ \mathrm{cm}^{-1}$.

Chl pair	Min	5%	50%	95%	Max
*a*602–*a*603	1.20	1.23	1.55	1.80	1.90
*a*602–*a*610	1.33	1.48	1.60	1.72	1.74
*a*602–*a*612	1.46	1.49	1.57	1.69	1.70
*a*602–*b*609	1.14	1.21	1.34	1.43	1.43
*a*603–*a*610	1.45	1.48	1.54	1.63	1.69
*a*603–*b*608	1.39	1.40	1.47	1.54	1.55
*a*603–*b*609	0.99	1.18	1.34	1.44	1.46
*a*610–*a*611	1.32	1.35	1.47	1.54	1.56
*a*610–*a*612	1.13	1.38	1.57	1.92	2.15
*a*611–*a*612	1.36	1.37	1.46	1.55	1.61
*b*608–*a*610	1.26	1.45	1.56	1.75	1.76
*b*608–*b*609	0.92	1.11	1.25	1.38	1.44

We note that V^{ρ}_{Coul} in environment displays a linear dependence on the vacuum coupling, irrespective of the Chl pair considered (figure 2(a)). This can be rationalized by representing the transition densities in terms of the corresponding transition dipoles (μ^{tr}) and quantifying their changes in terms of an enhancement and of a rotation of μ^{tr} in environment with respect to the vacuum. Indeed, we observe an enhancement of the transition dipole magnitude of 1.50 for Chls *a* and of 1.36 for Chls *b* on average (the distribution of the transition dipole magnitudes is shown in supplementary figure 1(a) (https://stacks.iop.org/JPCM/34/304004/mmedia)). On the contrary, the overall direction of the transition dipole in the plane of

the chlorophyll is not substantially affected by the environment, with 90% of the in-plane rotations lying in the interval $\left[-2.7^\circ, 0.6^\circ\right]$ for Chls *a* and in $\left[-4.8^\circ, 1^\circ\right]$ for Chls *b* (see supplementary figure 1(b) for the distribution of the in-plane rotations). The in-plane rotation of the transition dipole induced by the environment is slightly larger for Chls *b* than for Chls *a*. Overall, the transition dipoles of the chlorophylls are enhanced by the polarization, whereas their direction is almost unaltered and close to the NB–ND axis of the Chl ring as expected for the Q_y excitation considered here. This feature explains the almost linear dependence of V_{Coul} in environment on the same coupling in vacuum.

Let us now compare the Coulomb contribution to the electronic coupling V^{ρ}_{Coul} in environment and in vacuum with the total electronic coupling V^{ρ}_{Tot} (see figures 2(b) and (c), respectively). Neglecting the explicit contribution (like in figure 2(b)) results in large errors on the coupling value, up to 100 cm^{-1} for the largest couplings. On the other hand, estimating directly the coupling in vacuum results in smaller errors (figure 2(c)). Clearly, using directly the vacuum coupling benefits from a cancellation of errors. Overall, the coupling in vacuum is well correlated with the total coupling in environment (the Pearson's r^2 averaged over multiple Chl pairs is ~ 0.81).

This analysis tells us that the two environment effects, indirect and direct, affect the coupling in opposite ways. They both need to be included in the calculations in order to get a correct picture. By cancellation of errors, using vacuum couplings is better than considering only one of the two effects. Here, we analyzed only LHCII, but we expect this phenomenon to be quite general [30].

As already seen for the f^{ind} factors, the environment screens differently different Chl pairs, as shown in figure 2(b). The screening factor *s* (5) for the different Chl pairs is summarized in table 2. The screening effect depends on the proximity of paired Chls: *a*611–*a*612 and *a*603–*b*609 are the least screened, with *s* lying mostly between 0.73 and 0.84. However, there is no simple dependence on their distance, as the screening is also dependent on the specific environment experienced by each Chl [54–56]. For example, the pairs *a*602–*a*610, *a*602–*a*612, and *a*603–*a*610 experience a reduced screening despite the larger interpigment distances.

Table 2. Screening factors (5) for the selection of Chl pairs in LHCII. The fifth, 50th, and 95th quantiles of the f^{ind} distribution of each Chl–Chl pair are also reported. The factor is computed for points with $V^{\mathrm{env}}_{\mathrm{Coul}} > 10\ \mathrm{cm}^{-1}$.

Chl pair	Min	5%	50%	95%	Max
*a*602–*a*603	0.48	0.53	0.59	0.62	0.64
*a*602–*a*610	0.66	0.67	0.71	0.78	0.80
*a*602–*a*612	0.65	0.65	0.69	0.71	0.73
*a*602–*b*609	0.48	0.48	0.50	0.52	0.52
*a*603–*a*610	0.67	0.67	0.70	0.72	0.73
*a*603–*b*608	0.32	0.37	0.50	0.60	0.61
*a*603–*b*609	0.68	0.74	0.80	0.84	0.86
*a*610–*a*611	0.46	0.47	0.48	0.49	0.50
*a*610–*a*612	0.38	0.49	0.56	0.64	0.74
*a*611–*a*612	0.72	0.73	0.78	0.83	0.84
*b*608–*a*610	0.54	0.57	0.59	0.63	0.64
*b*608–*b*609	0.57	0.58	0.61	0.63	0.65

3.2. Validation of the TrEsp approximation in vacuum and in environment

In this section we validate the goodness of the TrEsp approximation (6) and (7) across a wide ensemble of Chl pairs and geometries. We have computed the TrEsp charges for multiple Chls along several frames of the LHCII MD. We are then able to discern whether the TrEsp charges can accurately account for fluctuations of the pigment geometries sampled in MD simulations.

The Coulomb coupling obtained from TrEsp charges $V^{\mathrm{TrEsp}}_{\mathrm{Coul}}$ (6) was compared with the Coulomb coupling V^{ρ}_{Coul} (1) computed from the transition densities ρ^{tr} of the chlorophyll pairs. The comparison is shown in figure 3(a), and clearly shows that the TrEsp approximation is almost perfect for a wide range of coupling values. Not only does the TrEsp approximation well represent average couplings among pairs, but it also well describes the coupling variations within each pair.

To test whether the TrEsp approximation remains valid when the composite environment (protein, cofactors, membrane and solvent) is included, we have repeated the comparison for the same chlorophylls and frames extracted from the LHCII MD, this time including the MMPol representation of the environment atoms. Namely, the environment was considered both in the transition density calculations and in the explicit contribution (7). This comparison is shown in figures 3(b) and (c). Figure 3(b) shows the screening contribution, as computed under the TrEsp–MMPol approximation $V^{\mathrm{TrEsp}}_{\mathrm{MMPol}}$ (7) and from the transition densities $V^{\rho}_{\mathrm{MMPol}}$ (2), denoting the excellent agreement between the explicit environment contribution computed in the two ways. Figure 3(c) shows the comparison between the total coupling (3) computed under the TrEsp–MMPol approximation $V^{\mathrm{TrEsp}}_{\mathrm{Tot}}$ and from the transition densities V^{ρ}_{Tot}. Clearly, the TrEsp approximation works equally well also when the environment is included.

The results reported in figure 3 show that the set of TrEsp charges for a particular Chl geometry is completely equivalent to the full transition density ρ^{tr}, from the point of view of the Coulomb coupling calculation. This holds true also for the couplings computed in the environment.

This framework allows simplifying the coupling calculations, but the transition densities still need to be computed for each pigment geometry. Indeed, we stress that the excellent agreement herein reported holds for TrEsp charges fitted at each particular Chl conformation. Moreover, the environment effect on the transition effect on the transition density has to be computed. In the following, we attempt to find the best possible approximation that allows us to estimate TrEsp charges without repeating QM/MM(Pol) calculations.

3.3. A regression approach for computing the TrEsp charges

When investigating fluctuations of the TrEsp couplings along an MD simulation, the commonly employed strategy is to fit the TrEsp charges once for an optimized geometry of the

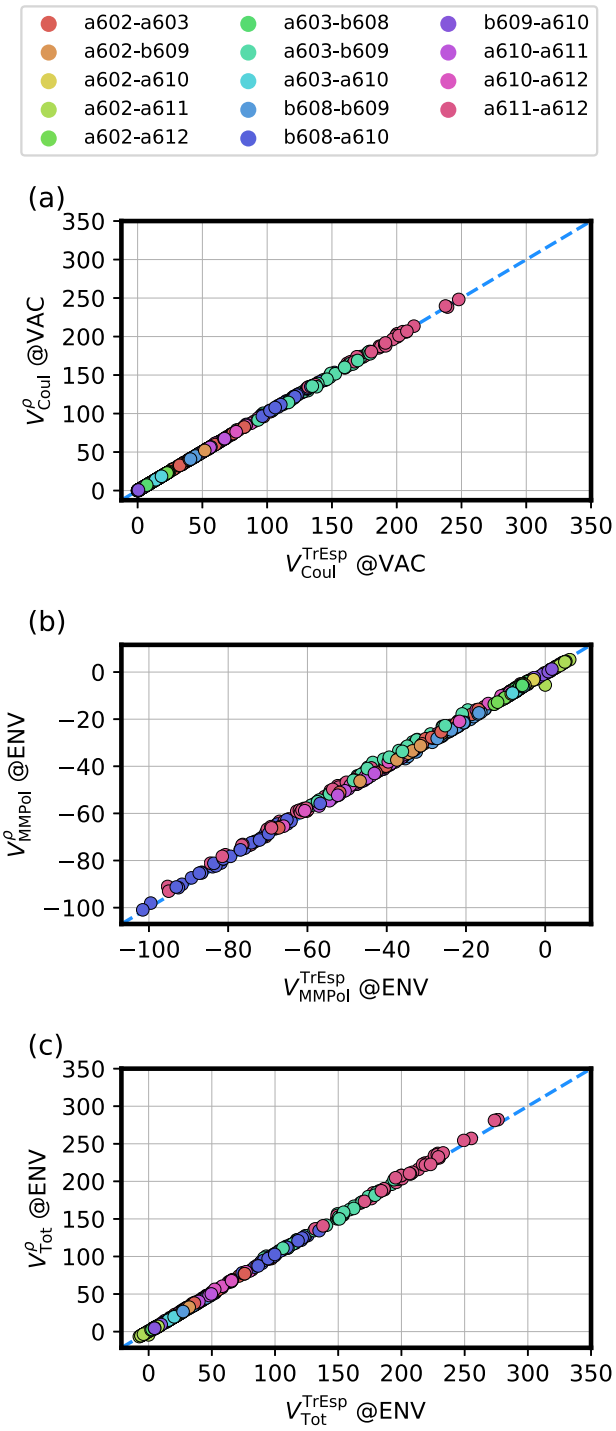

Figure 3. Performance of the TrEsp approximation in computing the various parts of the electronic coupling V for different chlorophyll pairs (top legend)(a) Coulomb contribution; (b) total coupling; (c) MMPol contribution. All values are in cm^{-1}.

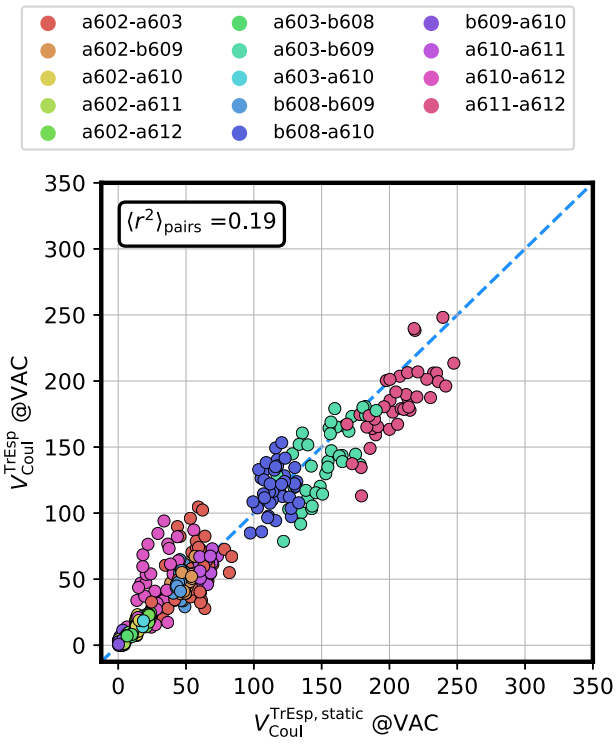

Figure 4. Performance of the static TrEsp approximation in computing the Coulomb coupling V_{Coul} in vacuum (@VAC). The Coulomb coupling computed from TrEsp charges obtained at the optimized geometry of either Chl *a* or Chl *b* is denoted as $V^{\mathrm{TrEsp,static}}_{\mathrm{Coul}}$, while the Coulomb coupling computed from TrEsp charges fitted for each geometry is denoted as $V^{\mathrm{TrEsp}}_{\mathrm{Coul}}$. The inset reports the Pearson's correlation coefficient squared r^2, computed for each Chl pair and averaged over the pairs, $\langle\cdot\rangle_{\mathrm{pairs}}$. Several Chl pairs have been considered (top legend). All values are in cm^{-1}.

chlorophyll, and then employ the fitted charges without further modifications along the MD [17–23]. This static-TrEsp approach is clearly a very cost-effective strategy, but it can introduce errors in the computation of the couplings.

In order to assess this error, we compare in figure 4 the static-TrEsp approximation ($V^{\mathrm{TrEsp,static}}_{\mathrm{Coul}}$) with the coupling computed from geometry-specific TrEsp charges ($V^{\mathrm{TrEsp}}_{\mathrm{Coul}}$).

The static TrEsp approach does distinguish different Chl pairs, because each pair has a specific intermolecular arrangement. However, within pairs, the static TrEsp shows evident limitations: Pearson's r^2 averaged over multiple Chl pairs is ~ 0.2, and the best fitted pair has r^2 around 0.55. We conclude that static-TrEsp approach is not suitable to investigate coupling fluctuations along a MD trajectory, and the geometry dependence of the TrEsp charges has to be taken into account.

As an alternative and efficient method, here we adopt a regression approach which aims to learn a mapping from the molecular geometry to the set of TrEsp charges. We have employed a ridge regression model (section 2.2), which we have found beneficial as opposed to using ordinary least-squares (OLS) regression. Although KRR is a more general model, we found that it provides no practical advantages over ridge (see supplementary figures 2 and 3), while requiring considerably more training time to select the kernel and regularization parameters. The success of a linear model probably reflects the intrinsic linear dependence of the transition density on the internal coordinates of the molecule.

Figure 5 shows that our regression yields good predictions of the Coulomb coupling across all pairs. The r^2 averaged over multiple Chl–Chl pairs is $\sim$0.95 (figure 5(a)). Furthermore, the accuracy of the predicted TrEsp charges is balanced across the Chl geometry: as shown in figure 5(b), the r^2 of the TrEsp

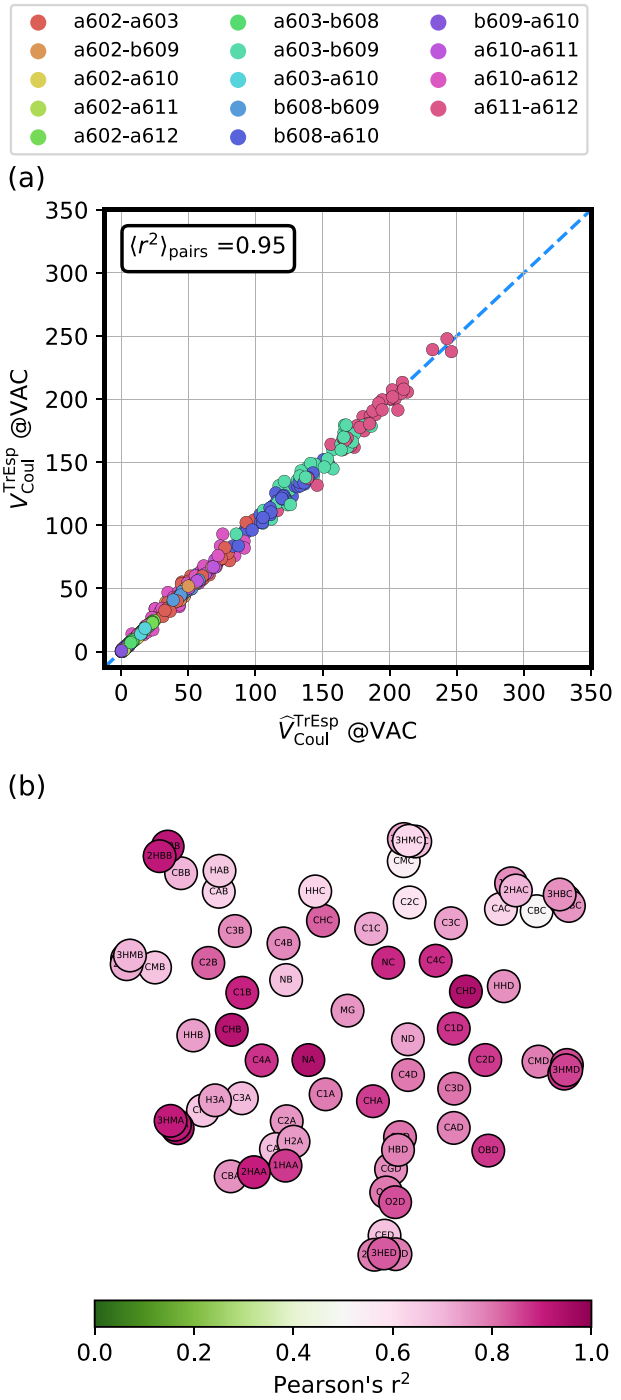

Figure 5. Performance of linear ridge regression with the permuted CM descriptor on the TrEsp charges and on the estimation of the Coulomb coupling V_{Coul}. (a) Comparison of the Coulomb coupling $\widehat{V}^{\text{TrEsp}}_{\text{Coul}}$ computed from predicted TrEsp charges and the Coulomb coupling $V^{\text{TrEsp}}_{\text{Coul}}$ computed from TrEsp charges (all values are in cm^{-1}). The inset reports the Pearson's correlation coefficient squared r^2, obtained for each Chl pair and averaged over the pairs, $\langle\cdot\rangle_{\text{pairs}}$. Several Chl pairs have been considered (top legend). (b) Plot of the Pearson's r^2 score over each atom of Chl *a* for a hold-out set. The r^2 score is computed between the set of predicted charges and the set of target charges. The bidimensional representation of the chlorophyll has been computed with multidimensional scaling over a chlorophyll's Cartesian distance matrix.

Table 3. Posterior mean and 95% highest posterior density interval (HPDI) for the parameters γ_m and β_{mn}, derived from a Bayesian linear model relating the Coulomb coupling V_{Coul} in environment to the one in vacuum. The model is detailed in the supplementary information.

Parameter	Mean	HPDI (low)	HPDI (up)
γ_a	1.21	1.19	1.24
γ_b	1.08	1.07	1.10
β_{aa}	1.47	1.41	1.53
β_{ab}	1.32	1.27	1.36
β_{bb}	1.18	1.15	1.21

charges on each atom is at least 0.6, with most atoms showing $r^2 > 0.8$.

The regression approach outlined here bypasses completely the need of performing QM calculations, and allows computing the Coulomb coupling accurately over an arbitrary number of MD frames. Moreover, the simplicity of the linear model allows us to obtain the regression parameters easily and with relatively few training samples (supplementary figure 2 shows the error as a function of the number of training samples).

3.4. Fast approximation of the coupling in environment

We have seen in section 3.1 that the indirect effect of the environment is to amplify the transition density in vacuum. Therefore, it is possible to employ the vacuum TrEsp charges to estimate the TrEsp charges in the environment with a suitable scaling factor. Assuming a different amplification for the transition density of Chls *a* and Chls *b*, we can express the Coulomb coupling in environment as $V^{\text{env}}_{\text{Coul},mn} = \beta_{mn} V^{\text{vac}}_{\text{Coul}}$, with m, n indicating the chlorophylls involved in the interaction. The slope $\beta_{mn} = \beta_{nm} = \gamma_m \gamma_n$ is assumed to be a function of the chlorophyll type of m and n, i.e., whether they are both Chls *a*, both Chls *b*, or one Chl *a* and one *b*.

We estimated the parameters γ_a and γ_b through a Bayesian linear model (see supplementary information for details), and their posterior mean and 95% density interval are reported in table 3. We note that β_{aa} and β_{bb} are well compatible with the enhancement observed for the transition dipole magnitude in Chls *a* and *b*. The small rotations of the transition dipoles, however, are included implicitly in the Bayesian model, resulting in coefficients that are slightly smaller than the transition dipoles enhancements. The effectiveness of approximating V_{Coul} in environment via a rescaling of the vacuum coupling is shown in supplementary figure 4.

We further note that this approach is similar to the calibrated Poisson-TrEsp model described in [30]. There, one Chl *a* pair in the water-soluble chlorophyll-binding protein (WSCP) was employed as an example. Calculations were performed after optimizing the Chl *a* coordinates extracted from the crystal structure, and the scaling factor was estimated by comparing vacuum and polarizable continuum model calculations. Here, the scaling factor β_{mn} is extracted from a Bayesian model that considers different pigment/environment configurations in dif-

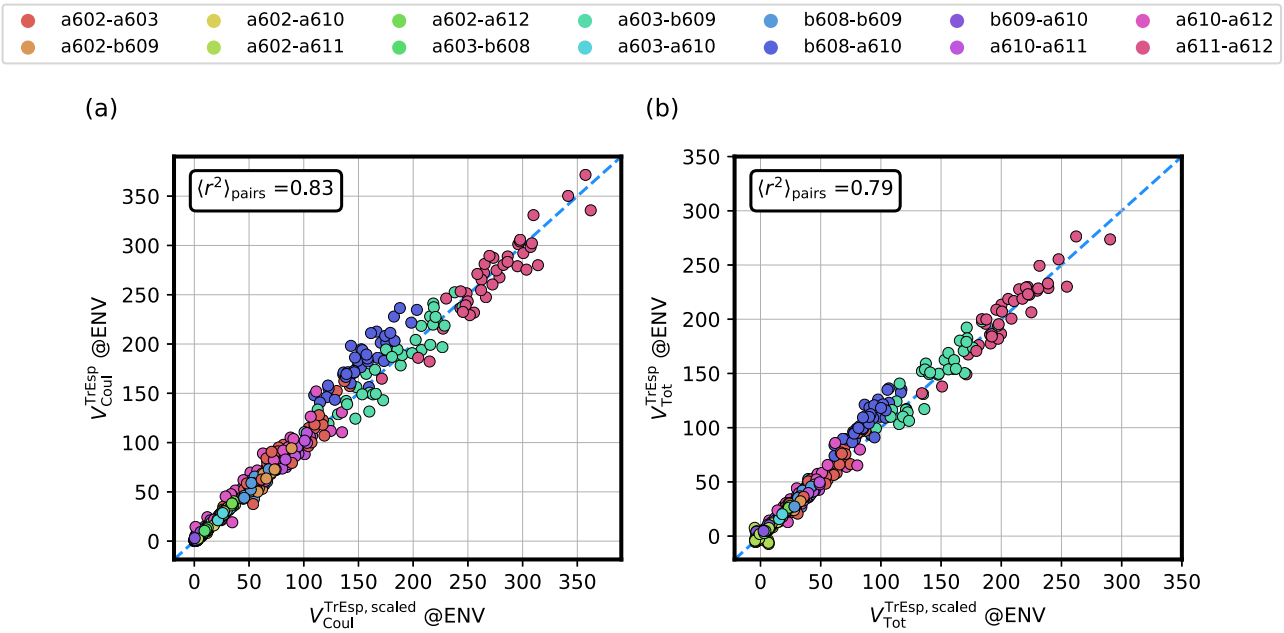

Figure 6. Approximating the TrEsp charges in environment through a rescaling of the predicted TrEsp charges in vacuum. (a) Comparison between the Coulomb coupling obtained from the TrEsp charges predicted by the linear ridge model and scaled by the γ_m factors, $V_{\text{Coul}}^{\text{TrEsp,scaled}}$, and the Coulomb coupling computed from TrEsp charges fitted for each Chl in each frame, $V_{\text{Coul}}^{\text{TrEsp}}$. (b) Comparison between the total coupling obtained from the predicted and scaled TrEsp charges $V_{\text{Tot}}^{\text{TrEsp,scaled}}$ and the total coupling from TrEsp charges fitted for each Chl and each frame $V_{\text{Tot}}^{\text{TrEsp}}$. The inset reports the Pearson's correlation coefficient squared r^2, computed for each Chl–Chl pair and then averaged over all the pairs, $\langle\cdot\rangle_{\text{pairs}}$.

ferent binding sites, and different relative arrangements of Chl pairs. The more diversified landscape of Chl pairs in LHCII as compared to WSCP adds to the robustness of the rescaling procedure, as demonstrated by supplementary figure 4, but also reveals its limitations. In fact, the specific polarization experienced by each Chl can only be captured partially by the present rescaling, see also supplementary table 1 and supplementary figure 5.

The rescaling approximation here proposed can be coupled with the regression model discussed in section 3.3. Indeed, by rescaling the predicted vacuum TrEsp charges with the γ_a, γ_b factors (table 3), it is possible to obtain the Coulomb coupling in environment with no additional costs. Furthermore, the explicit contribution of the environment can be computed from these rescaled charges using (7), recovering the full coupling. This is shown in figure 6, which evidences a fairly good agreement between the predicted coupling in environment and the same coupling computed from true TrEsp charges. The remarkable feature of this approach is that a model trained only on vacuum samples can be employed to compute the couplings in environment with a small error.

We further note that, given the correlation of the vacuum coupling with the total one analyzed in section 3.1 (figure 2(c)), it is also possible to use directly the predicted TrEsp charges in vacuum to compute the total environment coupling (the averaged Pearson's r^2 is ~ 0.76). We stress however that this is a particular case, and that computing explicitly the environment contribution as shown here is a more general and robust approach, which is able to recover the environment contributions in a transparent way.

This approach is extremely fast as the TrEsp charges prediction is on the order of ms, and employing the TrEsp charges to compute V_{MMPol} takes a couple of seconds, resulting in an instantaneous estimation of the coupling both in vacuum and in environment. This is to be compared with the time required to compute the transition density by means of a direct QM method for several Chls. Taking 10 min as a rough estimate to obtain the transition density for a single Chl, computing the transition densities for the 42 Chls in the three monomers of LHCII requires ~ 7 hours for a single MD frame. This demonstrates that approximating the couplings with our linear model is an accurate and orders of magnitude faster means of computing the electronic coupling between Chls in LH systems.

4. Conclusions

In this work, we have presented a fast method for computing the electronic couplings in multichromophoric systems by extending the transition charge approach (TrEsp) in two independent but parallel directions.

First, we have combined the TrEsp with a fully polarizable (MMPol) description of the environment. We have demonstrated that our approach, which we have named TrEsp–MMPol, is completely equivalent to a full QM/MMPol description based on transition densities and that this equivalence holds across all the inter-pigment distances and orientations sampled in our MD simulation.

To further enhance the computational efficiency of the TrEsp–MMPol approximation we have devised a regression

approach to predict the TrEsp charges directly from the chromophore geometry, i.e., bypassing completely the QM calculations. We have shown that the regression approach yields couplings with an accuracy that closely matches the reference QM one. Finally, we have leveraged the properties of the chlorophyll transition densities to devise a scaling procedure that allows approximating the environment coupling from the predicted vacuum charges.

In the present work we have sampled geometries from classical trajectories, but the method presented herein also applies to *ab initio* or semiempirical MD sampling. Such more sophisticated sampling would show how accurate TrEsp fluctuations are when obtained along classical MD trajectories.

Our method is simple, fast, and accurate, and opens the way to computing couplings accurately for a large number of geometries along MD trajectories. We believe that this method can contribute significantly when investigating the impact of protein dynamics on the pigment–pigment excitonic interactions. The prediction of vacuum and environment TrEsp charges along MD trajectories is implemented in a Python code available for download under the LGPL license agreement [57].

Acknowledgments

The authors acknowledge funding by the European Research Council, under the Grant ERC-AdG-786714 (LIFETimeS).

Data availability statement

The data that support the findings of this study are available upon request from the authors.

ORCID iDs

Edoardo Cignoni https://orcid.org/0000-0001-5392-8097
Lorenzo Cupellini https://orcid.org/0000-0003-0848-2908
Benedetta Mennucci https://orcid.org/0000-0002-4394-0129

References

[1] Mirkovic T, Ostroumov E E, Anna J M, van Grondelle R, Govindjee R and Scholes G D 2017 *Chem. Rev.* **117** 249–93
[2] Scholes G D, Fleming G R, Olaya-Castro A and van Grondelle R 2011 *Nat. Chem.* **3** 763–74
[3] Mikhnenko O V, Blom P W M and Nguyen T-Q 2015 *Energy Environ. Sci.* **8** 1867–88
[4] Collini E and Scholes G D 2009 *Science* **323** 369–73
[5] Chenu A and Scholes G D 2015 *Annu. Rev. Phys. Chem.* **66** 69–96
[6] Jang S J and Mennucci B 2018 *Rev. Mod. Phys.* **90** 35003
[7] Aragó J and Troisi A 2015 *Adv. Funct. Mater.* **26** 2316–25
[8] Segatta F, Cupellini L, Garavelli M and Mennucci B 2019 *Chem. Rev.* **119** 9361–80
[9] Jansen T L C 2021 *J. Chem. Phys.* **155** 170901
[10] You Z Q and Hsu C P 2014 *Int. J. Quantum Chem.* **114** 102–15
[11] Curutchet C and Mennucci B 2017 *Chem. Rev.* **117** 294–343
[12] Cupellini L, Corbella M, Mennucci B and Curutchet C 2019 *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **9** 1–23
[13] Madjet M E, Abdurahman A and Renger T 2006 *J. Phys. Chem.* B **110** 17268–81
[14] Kistler K A, Spano F C and Matsika S 2013 *J. Phys. Chem.* B **117** 2032–44
[15] Adolphs J, Müh F, Madjet M E-A and Renger T 2008 *Photosynth. Res.* **95** 197–209
[16] Renger T and Müh F 2012 *Photosynth. Res.* **111** 47–52
[17] Olbrich C, Jansen T L C, Liebers J, Aghtar M, Strümpfer J, Schulten K, Knoester J and Kleinekathöfer U 2011 *J. Phys. Chem.* B **115** 8609–21
[18] Van Der Vegte C P, Prajapati J D, Kleinekathöfer U, Knoester J and Jansen T L C 2015 *J. Phys. Chem.* B **119** 1302–13
[19] Daskalakis V, Maity S, Hart C L, Stergiannakos T, Duffy C D P and Kleinekathöfer U 2019 *J. Phys. Chem.* B **123** 9609–15
[20] Lapillo M, Cignoni E, Cupellini L and Mennucci B 2020 *Biochim. Biophys. Acta - Bioenerg.* **1861** 148282
[21] Daskalakis V, Papadatos S and Stergiannakos T 2020 *Chem. Commun.* **56** 11215
[22] Cignoni E, Lapillo M, Cupellini L, Acosta-Gutiérrez S, Gervasio F L and Mennucci B 2021 *Nat. Commun.* **12** 1–9
[23] Gray C, Wei T, Polívka T, Daskalakis V and Duffy C D P 2022 *Front. Plant Sci.* **12** 1–13
[24] Chrysafoudi A, Maity S, Kleinekathöfer U and Daskalakis V 2021 *J. Phys. Chem. Lett.* **12** 9626–33
[25] Curutchet C, Scholes G D, Mennucci B and Cammi R 2007 *J. Phys. Chem.* B **111** 13253–65
[26] Bondanza M, Nottoli M, Cupellini L, Lipparini F and Mennucci B 2020 *Phys. Chem. Chem. Phys.* **19** 14433–48
[27] Hsu C-P 2009 *Acc. Chem. Res.* **42** 509–18
[28] Iozzi M F, Mennucci B, Tomasi J and Cammi R 2004 *J. Chem. Phys.* **120** 7029–40
[29] Curutchet C, Muñoz-Losa A, Monti S, Kongsted J, Scholes G D and Mennucci B 2009 *J. Chem. Theory Comput.* **5** 1838–48
[30] Friedl C, Fedorov D G and Renger T 2022 *Phys. Chem. Chem. Phys.* **19** 5014–38
[31] Hsu C-P, Fleming G R, Head-Gordon M and Head-Gordon T 2001 *J. Chem. Phys.* **114** 3065
[32] Williams C K and Rasmussen C E 2006 *Gaussian Processes for Machine Learning* vol 2 (Cambridge, MA: MIT Press)
[33] Deringer V L, Bartók A P, Bernstein N, Wilkins D M, Ceriotti M and Csányi G 2021 *Chem. Rev.* **121** 10073–141
[34] Močkus J 1975 On Bayesian methods for seeking the extremum *Optimization Techniques IFIP Technical Conf.* (Springer) pp 400–4
[35] Jones D R, Schonlau M and Welch W J 1998 *J. Glob. Optim.* **13** 455–92
[36] Shahriari B, Swersky K, Wang Z, Adams R P and De Freitas N 2015 *Proc. IEEE* **104** 148–75
[37] Rupp M, Tkatchenko A, Müller K R and Von Lilienfeld O A 2012 *Phys. Rev. Lett.* **108** 1–5
[38] Hansen K, Montavon G, Biegler F, Fazli S, Rupp M, Scheffler M, Von Lilienfeld O A, Tkatchenko A and Müller K-R 2013 *J. Chem. Theory Comput.* **9** 3404–19
[39] Häse F, Valleau S, Pyzer-Knapp E and Aspuru-Guzik A 2016 *Chem. Sci.* **7** 5139–47
[40] Farahvash A, Lee C K, Sun Q, Shi L and Willard A P 2020 *J. Chem. Phys.* **153** 74111
[41] Krämer M, Dohmen P M, Xie W, Holub D, Christensen A S and Elstner M 2020 *J. Chem. Theory Comput.* **16** 4061–70
[42] Montavon G, Rupp M, Gobre V, Vazquez-Mayagoitia A, Hansen K, Tkatchenko A, Müller K R and Von Lilienfeld O A 2013 *New J. Phys.* **15** 95003

[43] Rupp M 2015 *Int. J. Quantum Chem.* **115** 1058–73
[44] Tchagang A B and Valdés J J 2019 Prediction of the atomization energy of molecules using Coulomb matrix and atomic composition in a Bayesian regularized neural networks *Int. Conf. on Artificial Neural Networks* (Springer) pp 793–803
[45] Balevičius V, Fox K F, Bricker W P, Jurinovich S, Prandi I G, Mennucci B and Duffy C D P 2017 *Sci. Rep.* **7** 1–10
[46] Cupellini L, Calvani D, Jacquemin D and Mennucci B 2020 *Nat. Commun.* **11** 662
[47] Wang J, Cieplak P, Li J, Hou T, Luo R and Duan Y 2011 *J. Phys. Chem.* B **115** 3091–9
[48] Frisch M J *et al* 2016 *Gaussian16 Revision C.01* Gaussian Inc. Wallingford CT
[49] Sláma V, Cupellini L and Mennucci B 2020 *Phys. Chem. Chem. Phys.* **22** 16783–95
[50] Guarnetti Prandi I, Sláma V, Pecorilla C, Cupellini L and Mennucci B 2022 *Commun. Biol.* **5** 1–10
[51] Wang C-I, Braza M K E, Claudio G C, Nellas R B and Hsu C-P 2019 *J. Phys. Chem.* A **123** 7792–802
[52] Wang C-I, Joanito I, Lan C-F and Hsu C-P 2020 *J. Chem. Phys.* **153** 214113
[53] Westermayr J and Marquetand P 2020 *J. Chem. Phys.* **153** 154112
[54] Curutchet C, Kongsted J, Muñoz-Losa A, Hossein-Nejad H, Scholes G D and Mennucci B 2011 *J. Am. Chem. Soc.* **133** 3078–84
[55] Jurinovich S, Viani L, Prandi I G, Renger T and Mennucci B 2015 *Phys. Chem. Chem. Phys.* **17** 14405–16
[56] Corbella M, Cupellini L, Lipparini F, Scholes G D and Curutchet C 2019 *ChemPhotoChem* **3** 945–56
[57] Cignoni E, Cupellini L, Lipparini F, Luongo C and Mennucci B 2022 EXCIPY: regression-based fast estimation of TrEsp charges and Coulomb coupling along all-atom molecular dynamics trajectories of light-harvesting complexes https://doi.org/10.5281/zenodo.6497263