Linear response time-dependent density-functional theory (LR-TDDFT)

Tim Zuehlsdorff

July 20, 2015

1 Linear Response TDDFT

The linear response TDDFT (LR-TDDFT) functionality in ONETEP allows the calculation of the low energy excited states of a system in linear scaling effort. In contrast to time-evolution TDDFT, where the density matrix of the system is propagated explicitly in time, LR-TDDFT recasts the problem of finding TDDFT excitation energies into an effective non-hermitian eigenvalue equation of the form:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \vec{\mathbf{X}} \\ \vec{\mathbf{Y}} \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \vec{\mathbf{X}} \\ \vec{\mathbf{Y}} \end{pmatrix} \tag{1}$$

where the elements of the block matrices $\mathbf{A}$ and $\mathbf{B}$ can be expressed in canonical Kohn-Sham representation as

$$A_{cv,c'v'} = \delta_{c,c'}\delta_{v,v'}(\epsilon_c^{\mathrm{KS}} - \epsilon_v^{\mathrm{KS}}) + K_{cv,c'v'} \tag{2}$$

$$B_{cv,c'v'} = K_{cv,v'c'} \tag{3}$$

Here, c and v denote Kohn-Sham conduction and valence states and $\mathbf{K}$ is the coupling matrix with elements given by

$$K_{cv,c'v'} = 2\int \mathrm{d}^3r\mathrm{d}^3r' \left[\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta^2 E_{\mathrm{xc}}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\bigg|_{\rho^{\{0\}}}\right] \times\psi_c^{\mathrm{KS}*}(\mathbf{r})\psi_v^{\mathrm{KS}}(\mathbf{r})\psi_{v'}^{\mathrm{KS}*}(\mathbf{r}')\psi_{c'}^{\mathrm{KS}}(\mathbf{r}'). \tag{4}$$

with E_{xc} being the exchange-correlation energy. Its second derivative, evaluated at the ground-state density $\rho^{\{0\}}$ of the system, is normally referred to as the exchange-correlation kernel.

The above equation can be understood as an effective 2-particle Hamiltonian consisting of a diagonal part of conduction-valence eigenvalue differences and a coupling term $K_{cv,c'v'}$ connecting individual Kohn-Sham excitations.

In ONETEP, LR-TDDFT is implemented both in terms of the full TDDFT eigenvalue equation (Eqn. 1) and in the Tamm-Dancoff approximation, a commonly used simplification to the full non-hermitian eigenvalue equation, where the off diagonal elements $\mathbf{B}$ are set to zero. The problem of calculating the TDDFT excitation energies thus becomes equivalent to solving the hermitian eigenvalue equation

$$\mathbf{A}\vec{\mathbf{X}} = \omega\vec{\mathbf{X}} \tag{5}$$

The Tamm-Dancoff approximation violates time-reversal symmetry and oscillator strength sum rules and can blue-shift strong peaks in the spectrum by up to 0.3 eV, however, dark states are typically left almost unaltered from their corresponding states in the Tamm-Dancoff approximation.

In the ONETEP code, the Tamm-Dancoff eigenvalue equation is re-expressed in terms of two sets of NGWFs, one optimised for the valence space (denoted as $\{\phi_\alpha\}$) and one optimised for a low energy subspace of the conduction manifold (denoted as $\{\chi_\beta\}$, see the documentation of the conduction NGWF optimiation functionality). Furthermore, the eigenvalue equation is solved iteratively for the lowest few eigenvalues using a conjugate gradient method. In order to do so we define the action $\mathbf{q}$ of operator $\mathbf{A}$ acting $\vec{\mathbf{X}}$ in conduction-valence NGWF space as

$$(q^{\chi\phi})^{\alpha\beta} = (P^{\{\mathrm{c}\}}H^{\chi}P^{\{1\}} - P^{\{1\}}H^{\phi}P^{\{\mathrm{v}\}})^{\alpha\beta} + (P^{\{\mathrm{c}\}}V_{\mathrm{SCF}}^{\{1\}\chi\phi}P^{\{\mathrm{v}\}})^{\alpha\beta}. \tag{6}$$

where $\mathbf{H}^{\chi}$ and $\mathbf{H}^{\phi}$ are the Hamiltonians in conduction and valence NGWF representation respectively, $\mathbf{P}^{\{c\}}$ and $\mathbf{P}^{\{v\}}$ denote the conduction and valence density matrices and $\mathbf{P}^{\{1\}}$ is the response density matrix, a representation of the trial vector $\vec{\mathbf{X}}$ in conduction-valence NGWF space. $V_{\mathrm{SCF}}^{\{1\}}$ is the first order response of the system due to the density $\rho^{\{1\}}(\mathbf{r})$ associated with $\mathbf{P}^{\{1\}}$. Under this redefinition of the action $\mathbf{A}$ in conduction-valence NGWF space, finding the lowest N_ω excitation energies is equivalent to minimising

$$\Omega = \sum_i^{N_\omega}\omega_i = \sum_i^{N_\omega}\left[\frac{\mathrm{Tr}\left[\mathbf{P}_i^{\{1\}\dagger}\mathbf{S}^{\chi}\mathbf{q}_i^{\chi\phi}\mathbf{S}^{\phi}\right]}{\mathrm{Tr}\left[\mathbf{P}_i^{\{1\}^{\dagger}}\mathbf{S}^{\chi}\mathbf{P}_i^{\{1\}}\mathbf{S}^{\phi}\right]}\right] \tag{7}$$

with respect to $\left\{\mathbf{P}_i^{\{1\}}\right\}$ under the constraint

$$\mathrm{Tr}\left[\mathbf{P}_i^{\{1\}\dagger}\mathbf{S}^{\chi}\mathbf{P}_j^{\{1\}}\mathbf{S}^{\phi}\right] = \delta_{ij}. \tag{8}$$

If all density matrices involved in the above expressions, ie. $\mathbf{P}^{\{1\}}$, $\mathbf{P}^{\{c\}}$ and $\mathbf{P}^{\{v\}}$ are truncated and thus become sparse, the algorithm scales as $O(N)$ with system size for a fixed number of excitation energies N_ω and as $O(N_\omega^2)$ with the number of excitation energies required.

A similar algorithm can be derived for the full TDDFT eigenvalue equation, where we make use of the change of variables $\mathbf{p} = \vec{\mathbf{X}} + \vec{\mathbf{Y}}$ and $\mathbf{q} = \vec{\mathbf{X}} - \vec{\mathbf{Y}}$. Each TDDFT excitation then has two effective density matrices, $\mathbf{P}^{\{p\}}$ and $\mathbf{P}^{\{q\}}$, associated with it that have the same structure as $\mathbf{P}^{\{1\}}$ in the Tamm-Dancoff approximation. The density matrices do obey an updated orthonormality constraint of the form

$$\frac{1}{2}\left(\mathrm{Tr}\left[\mathbf{P}_i^{\{p\}\dagger}\mathbf{S}^{\chi}\mathbf{P}_j^{\{q\}}\mathbf{S}^{\phi}\right] + \mathrm{Tr}\left[\mathbf{P}_i^{\{q\}\dagger}\mathbf{S}^{\chi}\mathbf{P}_j^{\{p\}}\mathbf{S}^{\phi}\right]\right) = \delta_{ij} \tag{9}$$

and an analogous expression for the total energy Ω in full TDDFT can be derived.

2 Performing a LR-TDDFT calculation

The LR-TDDFT calculation in ONETEP is enabled by setting the task flag to TASK=LR_TDDFT. The LR-TDDFT calculation mode reads in the density kernels and NGWFs of a converged ground state and conduction state calculation, so the .dkn, .dkn_cond, .tightbox_ngwfs and .tightbox_ngwfs_cond files all need to be present. The most important keywords in a TDDFT calculation are:

- `lr_tddft_RPA`: T/F.
  Boolean, default `lr_tddft_RPA`=F. If set to T, the code performs a full TDDFT calculation without relying on the simplified Tamm-Dancoff approximation.

- `lr_tddft_num_states`: n
  Integer, default `lr_tddft_num_states` = `1`.
  The keyword specifies how many excitations we want to converge. If set to a positive integer n, the TDDFT algorithm will converge the n lowest excitations of the system.

- `lr_tddft_cg_threshold`: x
  Real, default `lr_tddft_cg_threshold` = 10^{-6}.
  The keyword specifies the convergence tolerance on the sum of the n TDDFT excitation energies. If the sum of excitation energies changes by less than x in two consecutive iterations, the calculation is taken to be converged.

- `lr_tddft_maxit_cg`: n
  Integer, default `lr_tddft_maxit_cg` = `60`.
  The maximum number of conjugate gradient iterations the algorithm will perform.

- `lr_tddft_triplet`: T/F
  Boolean, default `lr_tddft_triplet` = `F`.
  Flag that decides whether the `lr_tddft_num_states` = `n` states to be converged are singlet or triplet states.

- `lr_tddft_write_kernels`: T/F
  Boolean, default `lr_tddft_write_kernels` = `T`.
  If the flag is set to T, the TDDFT response density kernels are printed out at every conjugate gradient iteration. These files are necessary to restart a LR_TDDFT calculation.

- `lr_tddft_restart`: T/F
  Boolean, default `lr_tddft_trestart` = `F`.
  If the flag is set to T, the algorithm reads in `lr_tddft_num_states` = `n` response density kernels in .dkn format and uses them as initial trial vectors for a restarted LR_TDDFT calculation.

- `lr_tddft_restart_from_TDA`: T/F
  Boolean, default `lr_tddft_trestart_from_TDA` = `F`.
  If the flag is set to T and `lr_tddft_RPA`: T, the code will read in already converged density kernels $\left\{\mathbf{P}_i^{\{1\}}\right\}$ and use them as a starting guess for a full TDDFT calculation such that $\mathbf{P}_i^{\{p\}} = \mathbf{P}_i^{\{q\}} = \mathbf{P}^{\{1\}}$. In many cases, the full TDDFT results are similar to the Tamm-Dancoff results and this strategy of starting the full TDDFT calculation leads to a rapid convergence.

- `lr_tddft_kernel_cutoff`: x
  Real, default `lr_tddft_kernel_cutoff` = `1000a`$_0$.
  Keyword sets a truncation radius on all response density kernels in order to achieve linear scaling computational effort with system size.

While the LR_TDDFT calculation can be made to scale linearly for a fixed number of excitations converged, it should be kept in mind that the algorithm needs to perform orthonormalisation procedures and thus scales as $O(N^2)$ with `lr_tddft_num_states`.

3 Truncation of the Response density matrix

To run a fully linear scaling TDDFT calculation the response density matrix has to be truncated by setting `lr_tddft_kernel_cutoff`. This truncation introduces numerical errors into the calculation, which mainly manifest themselves in the form that the response density matrices do no longer exactly obey a first order idempotency constraint that is placed on them. The idempotency constraint can be written in form of an invariance equation:

$$\mathbf{P}^{\{1\}'} = \mathbf{P}^{\{c\}}\mathbf{S}^{\chi}\mathbf{P}^{\{1\}}\mathbf{S}^{\phi}\mathbf{P}^{\{v\}} = \mathbf{P}^{\{1\}} \tag{10}$$

To measure the degree to which the invariance relation is violated we make use of a penalty functional $Q\left[\mathbf{P}^{\{1\}}\right]$ given by:

$$Q\left[\mathbf{P}^{\{1\}}\right] = \mathrm{Tr}\left[\left(\mathbf{P}^{\{1\}\dagger}\mathbf{S}^{\chi}\mathbf{P}^{\{1\}}\mathbf{S}^{\phi} - \mathbf{P}^{\{1\}'\dagger}\mathbf{S}^{\chi}\mathbf{P}^{\{1\}'}\mathbf{S}^{\phi}\right)^2\right]. \tag{11}$$

For truncated $\mathbf{P}^{\{1\}}$, $Q\left[\mathbf{P}^{\{1\}}\right] \neq 0$ which can lead to problems in the convergence of the conjugate gradient algorithm. In order to avoid these issues, the TDDFT routines perform the minimisation of the energy in an analogous form to the LNV method in ground-state calculations: The auxiliary density kernel $\mathbf{P}^{\{1\}'}$ is used instead of $\mathbf{P}^{\{1\}}$ for the minimisation of Ω. While $\mathbf{P}^{\{1\}'}$ is much less sparse than $\mathbf{P}^{\{1\}}$ it preserves idempotency to the same degree as the conduction and valence density kernel, yielding a stabilised convergence.

However, should $Q\left[\mathbf{P}^{\{1\}}\right]$ diverge significantly from 0 during the calculation, there are routines in place similar to the kernel purification schemes in ground state DFT that force the kernel towards obeying its idempotency constraint. The keyword controlling these routines are given below:

- `lr_tddft_penalty_tol`: x
  Real, default `lr_tddft_penalty_tol` = 10^{-8}.
  Keyword sets a tolerance for the penalty functional. If $Q\left[\mathbf{P}^{\{1\}}\right]$ is larger than `lr_tddft_penalty_tol` the algorithm will perform purification iterations in order to decrease the penalty value and force $\mathbf{P}^{\{1\}}$ towards the correct idempotency behaviour.

- `lr_tddft_maxit_pen`: n
  Integer, default `lr_tddft_maxit_pen` = 20.
  The maximum number purification iterations performed per conjugate gradient step.

4 More advanced TDDFT kernel truncation schemes

There are many situations where physical intuition allows one to specify a more sophisticated sparsity pattern than a uniform spherical kernel cutoff on $\mathbf{P}^{\{1\}}$ (or $\mathbf{P}^{\{p\}}$ and $\mathbf{P}^{\{q\}}$ for full TDDFT). For example, in pigment-protein complexes the excitations of interest retain a relative localisation on the pigment and one would ideally converge these states directly, without obtaining any spurious charge transfer states from the pigment to far away regions of the protein, that can arise due to failures in semi-local exchange correlation functionals. This can be achieved by introducing a new block into the input file of the form

%`block species_tddft_kernel`
label1 label 2 label3 ...
label5 ...
...
%`endblock species_tddft_kernel`

where the labels refer to atom labels. As an example, consider a pigment protein complex, where the pigment atoms are labelled H1, C1 etc. while the protein atoms are labelled H, C, etc. Then we can force the excitations of the system to be fully localised on the pigment by including

%`block species_tddft_kernel`
C1 H1 ...
%`endblock species_tddft_kernel`

This has the effect of setting all elements of $\mathbf{P}^{\{1\}}$ to zero that correspond to conduction or valence NGWFs centered on atoms of the environment. In this way the electrostatic effects of the environment are treated fully quantum mechanically, while no delocalisation into the protein is allowed. If one would like to introduce a coupling to the environment but wants to suppress any charge transfer coupling between the pigment and its environment, it is possible to specify

%`block species_tddft_kernel`
C1 H1 ...
C H ...
%`endblock species_tddft_kernel`

It is possible to specify an arbitrary number of subregions in the system in this way. It is also possible to list the same species in different lines, allowing for charge transfer interactions between some atom types of two regions but not others.

5 Preconditioning

The TDDFT eigenvalue problem is generally ill-conditioned, which can lead to a relatively slow convergence. For this reason, it is possible to precondition the eigenvalue problem, which is achieved by solving a linear system iteratively to a certain tolerance at each conjugate gradient step. Solving the linear system only requires matrix-matrix multiplications and is very cheap for small and medium sized systems, however, it can get more costly for very large systems, especially when no kernel truncation is used. In these cases, it can be necessary to reduce the number of default iterations of the preconditioner. The main keywords controlling the preconditioner are

- `lr_tddft_precond`: T/F
  Boolean, default `lr_tddft_precond` = `T`.
  Flag that decides whether the preconditioner is switched on or off.

- `lr_tddft_precond_iter`: n
  Integer, default `lr_tddft_precond_iter` = 20.
  Maximum number of iterations in the linear system solver applying the preconditioner.

- `lr_tddft_precond_tol`: x
  Real, default `lr_tddft_precond_tol` = 10^{-8}.
  The tolerance to which the linear system is solved in the preconditioner. Choosing a large tolerance means that the preconditioner is only applied approximately during each iteration.

6 Representation of the unoccupied subspace

In the LR_TDDFT method as implemented in ONETEP, the user has two options regarding the representation of the unoccupied subspace. The first option is to define the active unoccupied subspace of the calculation to only contain the Kohn-Sham states that were explicitly optimised in the COND calculation. The other is to

make use of a projector onto the entire unoccupied subspace, where we redefine the conduction density matrix as:

$$\mathbf{P}^{\{\mathrm{c}\}} = \left((\mathbf{S}^{\chi})^{-1} - (\mathbf{S}^{\chi})^{-1} \mathbf{S}^{\chi\phi} \mathbf{P}^{\{\mathrm{v}\}} \left(\mathbf{S}^{\chi\phi} \right)^{\dagger} (\mathbf{S}^{\chi})^{-1} \right). \tag{12}$$

The first option has the advantage that we only include states for which the NGWFs are well optimised, but has the drawback that some excitations converge very slowly with the size of the unoccupied subspace and thus a good convergence with the number of conduction states optimised is hard to reach. The second method implicitly includes the entire unoccupied subspace (to the extent that it is representable by a small, localised NGWF representation), but has the disadvantage that now states are included in the calculation for which the NGWFs are not optimised. Furthermore, the density matrix defined above is no longer strictly idempotent, leading to violations of the idempotency condition and thus a non-vanishing penalty functional $Q\left[\mathbf{P}^{\{1\}}\right]$, requiring kernel purification iterations as described in the previous section.

The problem of loss of idempotency can be avoided by using the joint NGWF set to represent the conduction space when using the projector. While this increases the computational cost of the LR_TDDFT calculation by a factor of 2, it preserves the idempotency of $\mathbf{P}^{\{\mathrm{c}\}}$ and is the recommended option when using the projector onto the unoccupied subspace.

The keywords controlling the use of the projector are

- `lr_tddft_projector`: T/F
  Boolean, default `lr_tddft_projector = T`.
  If the flag is set to T, the conduction density matrix $\mathbf{P}^{\{\mathrm{c}\}}$ is redefined to be a projector onto the entire unoccupied subspace.

- `lr_tddft_joint_set`: T/F
  Boolean, default `lr_tddft_joint_set = T`.
  If the flag is set to T, the joint NGWF set is used to represent the conduction space in the LR_TDDFT calculation.

7 Outputs

The LR_TDDFT calculation will produce a number of outputs. At the end of the calculation, the individual excitation energies and oscillator strengths will be computed and printed in the main ONETEP output file. Furthermore, the energies and oscillator strengths are used to generate a excitation spectrum written to the textfile

spectrum.dat. The peaks in the spectrum are Lorentzians whose width is defined by the radiative lifetimes of each excitation. Furthermore, by default, density cube files of the response density, the electron and the hole density for each excitation are printed out. The LR_TDDFT code can also perform an analysis of individual excitations, where the response density matrix is decomposed into dominant Kohn-Sham transitions. Since this analysis requires the Kohn-Sham eigenstates and thus a diagonalisation of the Hamiltonian, it scales as $O(N^3)$ and should not be performed for very large system sizes.

The keywords controlling these outputs are:

- `lr_tddft_write_densities`: T/F
  Boolean, default `lr_tddft_write_densities = T`.
  If the flag is set to T, the response density, electron density and hole density for each excitation is computed and written into a .cube file.

- `lr_tddft_analysis`: T/F
  Boolean, default `lr_tddft_analysis = F`.
  If the flag is set to T, a full $O(N^3)$ analysis of each TDDFT excitation is performed in which the response density is decomposed into dominant Kohn-Sham transitions.

8 Good practices and common problems

- The quality of the TDDFT excitation energies critically depends on the representation of the conduction space manifold. Any excitation that has a large contribution from an unoccupied state that is not explicitly optimised in the COND calculation is not expected to be represented correctly in the LR_TDDFT calculation. In general it is advisable to optimise as many conduction states as possible. However, high energy conduction states are often very delocalised and only representable if the conduction NGWF radius is increased significantly, thus leading to poor computational efficiency. In practice, there is a trade-off between computational efficiency and the representation of the conduction state manifold (see also the documentation on conduction state optimisation on this issue). Generally, TDDFT excitations should be converged with respect to both the conduction NGWF radius and the number of conduction states explicitly optimised.

- Since the ground state and conduction density kernels are used as projectors onto the occupied and unoccupied subspace in LR_TDDFT, one often finds

that the inner loop of the SINGLEPOINT and COND optimisation has to be converged to a higher degree of accuracy to achieve well behaved TDDFT results. It is therefore recommended to increase MAXIT_LNV and MINIT_LNV from their default value in the SINGLEPOINT and COND calculation. If no density kernel cutoff is used, the penalty functional value in the LR_TDDFT calculation should be vanishingly small. If the number increases significantly during a calculation or if the code begins to perform penalty optimisation steps, that is a clear sign that the initial conduction and valence density kernels are not converged well enough.

- In order to perform a LR_TDDFT calculation that scales fully linearly with system size, all density matrices involved have to be sparse and thus a KERNEL_CUTOFF has to be set for both the SINGLEPOINT and COND calculation. Using a density matrix truncation on the conduction states can sometimes be difficult depending on how the subspace of optimised conduction states is chosen and care has to be taken to prevent unphysical results.

- When running calculations in full linear scaling mode, the ground state and conduction density kernels are no longer strictly idempotent, which means that the penalty functional in LR_TDDFT will no longer be strictly zero. The code might perform penalty functional optimisation steps to keep the idempotency error small. However, these idempotency corrections can cause the conjugate gradient algorithm to stagnate and can even cause the energy to increase. If this happens, it is an indication that the minimum energy and maximum level of convergence for this truncation of the density kernel has been reached.

- When placing a truncation onto the the response density kernels it should be kept in mind that this may cause the optimisation to miss certain low energy excitations completely. Very long range charge-transfer type excitations cannot be represented by a truncated response density kernel and will thus be missing from the spectrum of excitations converged. However, well localised excitations should be unaffected. In a similar way, if the TDDFT kernel is limited to a certain region, it should be checked whether increasing the region leads to a smooth convergence of the energy of the localised state within the region.

9 Reference

For further background regarding the theory behind the LR_TDDFT method in ONETEP, as well as a number of benchmark tests, see

- Linear-scaling time-dependent density-functional theory in the linear response formalism, T. J. Zuehlsdorff, N. D. M. Hine, J. S. Spencer, N. M. Harrison, D. J. Riley, and P. D. Haynes, J. Chem. Phys. **139**, 064104 (2013)