**The potential of Imogolite nanotubes as (co-)photocatalyst: a linear-scaling Density Functional Theory study**

**E Poli,1,† J D Elliott1,† L E Ratcliff2, L Andrinopoulos2, J Dziedzic3,4, N D M Hine5, A A Mostofi2, C-K Skylaris3, P D Haynes2 and G Teobaldi1**

1 Stephenson Institute for Renewable Energy and Department of Chemistry, University of Liverpool, Liverpool L69 3BX, UK

2 The Thomas Young Centre for Theory and Simulation of Materials, Imperial College London, London SW7 2AZ, UK

3 School of Chemistry, University of Southampton, Southampton SO17 1BJ, UK

4 Faculty of Applied Physics and Mathematics, Gdansk University of Technology, Gdansk, Poland

5 Department of Physics, University of Warwick, Coventry CV4 7AL, UK

† These authors contributed equally to the manuscript

E-mail: g.teobaldi@liv.ac.uk

**Abstract**

We report a linear-scaling Density Functional Theory (DFT) study of the structure, wall-polarization absolute band-alignment and optical absorption of several, recently synthesized, open-ended Imogolite (Imo) nanotubes (NTs), namely single-walled (SW) aluminosilicate (AlSi), SW aluminogermanate (AlGe) , SW methylated aluminosilicate (AlSi-Me), and double-walled (DW) AlGe NTs. Simulations with three different semi-local and dispersion-corrected DFT-functionals reveal that the NT wall-polarization can be increased by nearly a factor of four going from SW-AlSi-Me to DW-AlGe. Absolute vacuum alignment of the NT electronic bands and comparison with those of rutile and anatase TiO2 suggest that the NTs may exhibit marked propensity to both photo-reduction and hole-scavenging. Characterization of the NTs’ band-separation and optical properties reveal the occurrence of (near-)UV inside-outside charge-transfer excitations, which may be effective for electron-hole separation and enhanced photocatalytic activity. Finally, the effects of the NTs’ wall-polarization on the absolute alignment of electron and hole acceptor states of interacting water (H2O) molecules are quantified and discussed.

Keywords: Photocatalysis, Inorganic nanotubes, Imogolite nanotubes, Linear-scaling Density Functional Theory

**1. README for the data catalogue**

This archive contains open-data for "The potential of Imogoliteinanotubes as (co-)photocatalyst: a linear-scaling Density Functional Theory study" by Poli, Emiliano, Elliott, Joshua, Ratcliff, LE , Andrinopoulos, L , Dziedzic, J, Hine, NDM , Mostofi, AA , Skylaris, CK , Haynes, PD and Teobaldi, Gilberto. To be published in the Journal of Physics: Condensed Matter

The open data is organized as follows

==== Figure-\*.agr
xmgrace files and raw data of the figures in the paper and supporting information

===  Al\_Si\_24.dat
Example of ONETEP input file

Description of the Methods used to generate the data follows in Section 2.

**2. Methods**

*2.1 Linear-scaling DFT in ONETEP*

In this section, we briefly present the theoretical framework behind linear-scaling DFT (LS-DFT) and its implementation in the ONETEP program [121-124]. The interested reader is referred to [120] for a recent review of LS-DFT methods.

Linear-scaling methods make use of the “nearsightedness” [125,126] inherent in quantum many-body systems by exploiting the localization of Wannier functions [127-130] or the single-particle density matrix, ρ(**r**,**r’**) [131,132]. ONETEP is based on the latter approach and on a formulation of DFT theory with norm-conserving pseudopotentials [133]. Within ONETEP ρ(**r**,**r’**), is expressed in a separable form [134,135] via atom-centred functions (Non-Orthogonal Generalized Wannier Functions, NWGFs [84]), ϕα(**r**), as:

  (1)

where Kαβ are the matrix elements of the density kernel, which are nonzero only if |rα – rβ| < rc, where rα and rβ indicate the coordinates of the centres of ϕα and ϕβ, and rc is a real-space cut-off threshold. The (optional) truncation of the density kernel (Kαβ), leading to a sparse density matrix [ρ(**r**,**r’**)] is justified by the known exponential decay of ρ(**r**,**r’**) with respect to |**r** – **r’**| for systems with an electronic BG [130-132], which makes any insulating or semiconducting systems (including the Imo-NTs considered here) amenable to LS-DFT simulation.

The NGWFs are centred on the nuclear coordinates and strictly localized within a sphere of radius Rα. Being non-orthogonal, the NGWFs are characterized by a non-diagonal overlap matrix, Sαβ:

  (2)

The NGWFs are in turn expanded as a linear combination, of coefficients Cmα, of localized yet orthogonal periodic cardinal sine (*psinc*) functions [136], Dm(**r**), as:

 (3)

with m indexing the real-space Cartesian grid inside the spherical localisation region of ϕα. The *psinc* functions are formed from a discrete sum of plane-waves, which makes the set of Dm(**r**) independent of the nuclear coordinates and systematically improvable upon increase of the kinetic energy cutoff [133]. In the ONETEP approach, the total DFT energy is minimized self-consistently with respect to Kαβ and Cmα in two nested loops [121-124]. As a result, the NGWFs are optimized *in situ* by iteratively improving the set of coefficients Cmα that minimize the total energy under the constraints of idempotency of the density matrix (ρ2 = ρ) and conservation of the number of electrons in the simulated system. The approach has been shown to lead to convergence of results to a near complete basis set quality even for minimal number of NGWFs employed in the simulations [137].

An alternative approach to self-consistent energy minimization is to instead employ a single loop that optimizes the elements of Kαβ only, maintaining the NGWFs fixed. Recent additions to ONETEP allow generation of suitable multiple-zeta basis sets out of pseudoatomic orbitals (PAOs), which can also be used with high accuracy given a large enough basis and the explicit calculation of the Pulay forces arising from incomplete optimization of the basis set [138].

As discussed in [121-124], the convergence of the ONETEP approach can depend on interlinked computational factors such as the kinetic energy cutoff, the number of NGWFs (ϕα) per atom and their localization radius. The reader if referred to [114] for initial benchmarking of the procedure on Imo NTs.

*2.2 Optimization of empty Kohn-Sham (KS) states and Fermi golden rule optical spectra in ONETEP*

The nested loops optimization of the DFT energy, density matrix [ρ(**r**,**r’**)], density kernel (Kαβ) and NGWFs (ϕα) allows linear-scaling solution of the Kohn-Sham DFT problem for systems with a non-zero electronic BG without explicit reference to individual KS states as typical for standard plane-wave DFT. KS states can, however, be recovered from a single diagonalization of the Hamiltonian at the end of a converged calculation [139]. Although capable of accurate computation of occupied KS states, the approach nevertheless results in poor description of all but the lowest-lying unoccupied KS states since the NWGFs are optimized to describe only the occupied KS states, thence ρ(**r**,**r’**). As a result, the description of empty KS states in terms of both energy and real-space amplitude, obtained from the standard optimization of valence NGWFs, can be unsatisfactory, even omitting some higher energy (but still bound) empty KS states entirely [140].

To overcome this limitation, which prevents highly accurate calculation of KS DFT BGs, and extension towards time-dependent (TD-DDFT) or perturbative post-DFT GW approaches [141], ONETEP implements a procedure for the optimization of a second set of atom-centered conduction NGWFs (χα) to correctly compute empty bound KS from pre-optimized valence NGWFS (ϕα) and ensuing density matrix [ρ(**r**,**r’**)]. The approach, described in [140], is based on the optimization of a selected number of empty KS states via nested minimization of the energy of a conduction-projected Hamiltonian (Hχp):

 (4)

with respect to the conduction NGWFs (χα) and related density kernel (Mαβ). Hχp is obtained from valence NGWFs (ϕα), density kernel Kαβ and overlap matrix (Sϕ, Eq. 2):

  (5)

Hχ is the unprojected conduction Hamiltonian, Tαβ = ∫dr ϕα\*(**r**) χβ(**r**) is the cross-overlap between valence- and conduction-NGWFs, and σ is an energy shift applied to ensure the applicability of the procedure also to initially positive eigenvalues of the (unprojected) conduction Hamiltonian. In analogy with the optimization of the valence density Kernel (Kαβ), and especially for low-energy empty KS-states with some degree of localization (i.e. occurrence of a gap in the unoccupied KS manifold), Mαβ can be truncated via a strict distance based cutoff, allowing the procedure to scale linearly with the size of the modelled system.

Once the conduction NGWFs have been optimized, the KS-Hamiltonian can be computed and diagonalized in a joint basis encompassing both valence- (ϕα) and conduction- (χα) optimized NGWFs, yielding eigenvalues in very good agreement with the results from standard plane-wave DFT (for the occupied and optimized empty KS states). Besides, the definition of KS BGs, the optimized conduction KS states can additionally be used for calculation of optical spectra via Fermi’s golden rule [140] or within the linear-response time-dependent DFT (LR-TDDFT) framework recently implemented in ONETEP [141].

As explained in [140], and considered in [114] for the Imo NTs, the convergence of this procedure can be affected by several factors such as the energy and number of the optimized conduction KS states, the similarity between valence and conduction KS states, the number of conduction NGWFs (χα) per atom and their localization radius.

Owing to the limitation of the adiabatic local density approximation (ALDA) for LR-TDDFT simulation of periodic semiconductors and insulators [142], optical absorption spectra were approximated on the basis of the imaginary part of the dielectric function ε2(ω) as available in ONETEP [140]:

  (6)

where *v* and *c* denote valence and conduction KS-state, respectively. |ψn,**k**> is the nth KS-state, of energy En,k at the kth **k**-point. Ω is the simulation cell volume. **q** is the polarization direction of the incoming photon of energy ħω. **r** is the position operator. As explained in [140,143-145], owing to the ill-undefined nature of **r** in periodic systems, ε2(ω) needs to be computed via the momentum operator (**p**) and then related to the position (**r**) representation through corresponding commutator with the Hamiltonian. In all cases, optical spectra were computed for extended x2 supercells of the tubes based on KS-states at the Γ-point only.

Although limitations in the procedure are to be expected owing to the neglect of non-local electron-hole interactions and relaxations in the simulated spectra [142], the adopted procedure should serve as initial exploration of the role of the different composition and curvature for the optical properties of Imo NTs.

*2.3 Further computational details*

Simulation of the SW AlSi and SW/DW AlGe NTs were performed with the PBE functional [146]. Owing to the presence of highly polarisable methyl groups, and to test possible deficiencies of the PBE functional, SW AlSi-Me NTs were simulated with PBE, PBE with empirical dispersion corrections according to Grimme’s D2 parameterization [147] (PBE-D2) and via the self-consistent dispersion functional OPTB88, chosen for the reported optimal performance in simultaneous treatment of H-bonding and methyl-group dispersion interactions [148]. In all cases, separable (Kleinman-Bylander) norm-conserving pseudopotentials [149], constructed with the opium code [150], were used.

As in [114], the adopted kinetic energy cutoff was 1000 eV and 4 (9) valence and conduction NGWFs were used for O (Al,Si) atoms. 1 NGWF was used for H atoms. In all cases, no truncation of the density kernel (Kαβ) was enforced. The localization radius for the valence (conduction) NGWFs was 8 Bohr (15 Bohr). All simulations were performed with periodic boundary conditions ensuring at least 15 Å vacuum separation between replicated images along the non-periodic directions. The conduction NGWFs optimization was performed on x2 super-cells along the tube-axis (i.e. two-fold replicas of the original minimal simulation cell) to accommodate the extended localization radius, which, together with optimization of 100 empty Kohn-Sham states, was necessary for sub-meV convergence of BGs and CB edges. The periodic unit for the SW-AlSi, SW-AlSi-Me, SW-AlGe and DW-AlGe NTs was the PBE-optimized one [106,114] (SW-AlSi: 8.666 Å, SW-AlSi-Me: 8.666 Å, SW-AlGe: 8.627 Å, DW-AlGe: 8.627 Å).

Geometry-relaxations where performed via the quasi-Newton optimization scheme based on the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [151].

To prevent biases due to the asymmetric distribution of the H2O molecules inside and outside the NTs (end ensuing dipoles), vacuum-aligned KS-states were obtained from 1D-periodic simulations with truncated electrostatics in the two directions perpendicular to the tube axis [152]. The adopted truncation radii (SW-AlSi: 57.2 Å, SW-AlSi-Me: 77.9 Å, SW-AlGe: 76.2 Å, DW-AlGe: 102.7 Å) were numerically checked to yield sub-meV convergence of the results. All the simulations made use of the hybrid MPI-OpenMP parallelism available in ONETEP [153].

**References**

[121] Skylaris C K, Haynes P D, Mostofi A A, and Payne M C 2015 *J. Chem. Phys.* **122** 084119

[122] Haynes P D, Skylaris C K, Mostofi A A, and Payne M C 2006 *Phys. Status Solidi B* **243** 2489

[123] Hine N D M, Haynes P D, Mostofi A A, Skylaris K, and Payne M C 2009 *Comput. Phys. Commun.* **180** 1053.

[124] Hine N D M, Robinson M, Haynes P D, Skylaris C K, Payne M C and Mostofi A A 2011 *Phys. Rev. B.* **83** 195102.

[125] Kohn W 1996 *Phys. Rev. Lett.* **76** 3168

[126] Prodan E and Kohn W 2005 *Proc. Natl. Acad. Sci. USA* **102** 11635

[127] Kohn W 1959 *Phys. Rev.* **115** 809

[128] des Cloizeaux J 1964 *Phys. Rev.* **135** A698.

[129] Nencieu G 1983 *Commun.* *Math. Phys.* **91** 81

[130] He L and Vanderbilt D 2001 *Phys. Rev. Lett.* **86** 5341

[131] des Cloizeaux J 1964 *Phys. Rev.* **135** A685.

[132] Ismail-Beigi S and Arias T A 1999 *Phys. Rev. Lett.* **82** 2127

[133] Payne M C, Teter M, Allan D C, Arias T A, and Joannopoulos J D 1992 *Rev. Mod. Phys.* **64** 1045

[134] McWeeny R 1960 *Rev. Mod. Phys.* **32** 335

[135] Hernandez E and Gillan M J 1995 *Phys. Rev. B* **51** 10157

[136] Skylaris C K, Mostofi A A, Haynes P D, Dieguez O, and Payne M C 2002 *Phys. Rev. B* **66** 035119

[137] P. D. Haynes, C. K. Skylaris, A. A. Mostofi, and M. C. Payne: *Chem. Phys. Lett.*, 2006, **422**, (4-6), 345-349

[138] Ruiz-Serrano A, Hine N D M and Skylaris C K 2012 *J. Chem. Phys.* **136** 234101

[139] Skylaris K, Haynes D, Mostofi A A and Payne M C 2005 *J. Phys.: Condens. Matter* **17** 5757

[140] Ratcliff L E, Hine N D M and Haynes D 2011 *Phys. Rev. B.* 2011 **84** 165131

[141] Zuehlsdorff T J, Hine N D M, Spencer J S, Harrison N M, Riley D J, and Haynes P D 2013 *J. Chem. Phys.* **139** 064104

[142] Onida G, Reining L and Rubio A 2002 *Rev. Mod. Phys.* **74** 601

[143] Pickard C J Ph.D. Thesis University of Cambridge 1997

[144] Read A J and Needs R J 1991 *Phys. Rev. B* **44** 13071

[145] Motta C, Giantomassi M, Cazzaniga M, Gal-Nagy K, and Gonze X 2010 *Comput. Mater. Sci.* **50** 698

[146] Perdew J, Burke K, and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865

[147] Grimme S 2006 *J. Comput. Chem.* **27** 1787

[148] Klimeš J, Bowler D R, Michaelides A 2010 *Journal of Physics: Condensed Matter* **22** 022201

[149] Gonze X, Stumpf R, Scheffler M 1991 *Phys. Rev. B* **44** 8503

[150] http://opium.sourceforge.net/sci.html

[151] Pfrommer B G, Côté M, Louie S G, and Cohen M. L 1997 *J. Comput. Phys.* **131** 233

[152] Hine N D M, Dziedzic J, Haynes P. D, Skylaris C-K. 2011 *J. Chem. Phys.* **135** 204103

[153] Wilkinson K A, Hine N D M, and Skylaris C-K 2014 *J. Chem. Theory Comput.* **10** 4782