

# Supporting Information: The OpenMolcas *Web*: A Community-Driven Approach to Advancing Computational Chemistry

Giovanni Li Manni<sup>1,65</sup>, Ignacio Fdez. Galván<sup>\*,2,65</sup>, Ali Alavi<sup>1,3</sup>, Flavia Aleotti<sup>4</sup>, Francesco Aquilante<sup>5</sup>, Jochen Autschbach<sup>6</sup>, Davide Avagliano<sup>4</sup>, Alberto Baiardi<sup>7</sup>, Jie J. Bao<sup>8</sup>, Stefano Battaglia<sup>2,9</sup>, Letitia Birnoschi<sup>10</sup>, Alejandro Blanco-González<sup>11</sup>, Sergey I. Bokarev<sup>12,13</sup>, Ria Broer<sup>14</sup>, Roberto Cacciari<sup>15</sup>, Paul B. Calio<sup>16</sup>, Rebecca K. Carlson<sup>8</sup>, Rafael Carvalho Couto<sup>17</sup>, Luis Cerdán<sup>18,19</sup>, Liviu F. Chibotaru<sup>20</sup>, Nicholas F. Chilton<sup>10</sup>, Jonathan Richard Church<sup>21</sup>, Irene Conti<sup>4</sup>, Sonia Coriani<sup>22</sup>, Juliana Cuéllar-Zuquin<sup>18</sup>, Razan E. Daoud<sup>15</sup>, Nike Dattani<sup>23,24</sup>, Piero Decleva<sup>25</sup>, Coen de Graaf<sup>26,27</sup>, Mickaël G. Delcey<sup>17</sup>, Luca De Vico<sup>15</sup>, Werner Dobrautz<sup>28</sup>, Sijia S. Dong<sup>8,29</sup>, Rulin Feng<sup>6,30</sup>, Nicolas Ferré<sup>31</sup>, Michael Filatov(Gulak)<sup>32</sup>, Laura Gagliardi<sup>16,8</sup>, Marco Garavelli<sup>4</sup>, Leticia González<sup>33</sup>, Yafu Guan<sup>34</sup>, Meiyuan Guo<sup>35</sup>, Matthew R. Hennefarth<sup>16</sup>, Matthew R. Hermes<sup>16,8</sup>, Chad E. Hoyer<sup>8,36</sup>, Miquel Huix-Rotllant<sup>31</sup>, Vishal Kumar Jaiswal<sup>4</sup>, Andy Kaiser<sup>13</sup>, Danil S. Kaliakin<sup>11</sup>, Marjan Khamesian<sup>2</sup>, Daniel S. King<sup>16</sup>, Vladislav Kochetov<sup>37</sup>, Marek Krośnicki<sup>38</sup>, Arpit Arun Kumar<sup>23</sup>, Ernst D. Larsson<sup>39</sup>, Susi Lehtola<sup>40,41</sup>, Marie-Bernadette Lepetit<sup>42,43</sup>, Hans Lischka<sup>44</sup>, Pablo López Ríos<sup>1</sup>, Marcus Lundberg<sup>45</sup>, Dongxia Ma<sup>1,8</sup>, Sebastian Mai<sup>33</sup>, Philipp Marquetand<sup>33</sup>, Isabella C. D. Merritt<sup>46</sup>, Francesco Montorsi<sup>4</sup>, Maximilian Mörchen<sup>7</sup>, Artur Nenov<sup>4</sup>, Vu Ha Anh Nguyen<sup>47,48</sup>, Yoshio Nishimoto<sup>49</sup>, Meagan S. Oakley<sup>8</sup>, Massimo Olivucci<sup>15,11</sup>, Markus Oppel<sup>33</sup>, Daniele Padula<sup>15</sup>, Riddhish Pandharkar<sup>16,8</sup>, Quan Manh Phung<sup>50,51</sup>, Felix Plasser<sup>52</sup>, Gerardo Raggi<sup>2,53</sup>, Elisa Rebolini<sup>54</sup>, Markus Reiher<sup>7</sup>, Ivan Rivalta<sup>4,55</sup>, Daniel Roca-Sanjuán<sup>18</sup>, Thies Romig<sup>37</sup>, Arta Anushirwan Safari<sup>1</sup>, Aitor Sánchez-Mansilla<sup>26</sup>, Andrew M. Sand<sup>8,56</sup>, Igor Schapiro<sup>21</sup>, Thais R. Scott<sup>8,16,57</sup>, Javier Segarra-Martí<sup>18</sup>, Francesco Segatta<sup>4</sup>, Dumitru-Claudiu Sergentu<sup>6,58</sup>, Prachi Sharma<sup>8</sup>, Ron Shepard<sup>59</sup>, Yanan Shu<sup>8</sup>, Jakob K. Staab<sup>10</sup>, Tjerk P. Straatsma<sup>60,61</sup>, Lasse Kragh Sørensen<sup>62</sup>, Bruno Nunes Cabral Tenorio<sup>22</sup>, Donald G. Truhlar<sup>8</sup>, Liviu Ungur<sup>47</sup>, Morgane Vacher<sup>46</sup>, Valera Veryazov<sup>39</sup>, Torben Arne Voß<sup>13</sup>

Oskar Weser<sup>1</sup>, Dihua Wu<sup>8</sup>, Xuchun Yang<sup>11</sup>, David Yarkony<sup>63</sup>, Chen  
Zhou<sup>8</sup>, J. Patrick Zobel<sup>33</sup> and Roland Lindh<sup>\*,2,64,65</sup>

- 1 *Electronic Structure Theory Department, Max Planck Institute for Solid State Research, Heisenbergstraße 1, 70569 Stuttgart, Germany*
- 2 *Department of Chemistry – BMC, Uppsala University, P. O. Box 576, SE-75123 Uppsala, Sweden*
- 3 *Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*
- 4 *Department of Industrial Chemistry "Toso Montanari", University of Bologna, 40136 Bologna, Italy*
- 5 *Theory and Simulation of Materials (THEOS) and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*
- 6 *Department of Chemistry, University at Buffalo, State University of New York, Buffalo, NY 14260-3000, USA*
- 7 *ETH Zurich, Laboratory for Physical Chemistry, Vladimir-Prelog-Weg 2, 8093 Zurich, Switzerland*
- 8 *Department of Chemistry, Chemical Theory Center, and Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431, United States*
- 9 *Current affiliation: Department of Chemistry, University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland*
- 10 *The Department of Chemistry, The University of Manchester, M13 9PL, Manchester, UK*
- 11 *Chemistry Department, Bowling Green State University, Overmann Hall, Bowling Green, OH 43403, USA*
- 12 *Chemistry Department, School of Natural Sciences, Technical University of Munich, Lichtenbergstr. 4, 85748 Garching, Germany*
- 13 *Institut für Physik, Universität Rostock, Albert-Einstein-Str. 23-24, 18059 Rostock, Germany*
- 14 *Theoretical Chemistry, Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands*
- 15 *Dipartimento di Biotecnologie, Chimica e Farmacia, Università di Siena, Via A. Moro 2, 53100 Siena, Italy*
- 16 *Department of Chemistry, Pritzker School of Molecular Engineering, James Franck Institute, Chicago Center for Theoretical Chemistry, The University of Chicago, Chicago, Illinois 60637, United States*
- 17 *Division of Theoretical Chemistry and Biology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, SE-106 91 Stockholm, Sweden*
- 18 *Instituto de Ciencia Molecular, Universitat de València, Catedrático José Beltrán Martínez n.º 2, 46980 Paterna, Spain*
- 19 *Instituto de Óptica (IO-CSIC), Consejo Superior de Investigaciones Científicas, 28006, Madrid, Spain*
- 20 *Department of Chemistry, KU Leuven, Celestijnenlaan 200F, 3001 Leuven, Belgium*
- 21 *Institute of Chemistry, The Hebrew University of Jerusalem*
- 22 *Department of Chemistry, Technical University of Denmark, Kemitorvet Bldg 207, 2800 Kongens Lyngby, Denmark*

- 23 HPQC Labs, Waterloo, N2T 2K9 Ontario, Canada
- 24 HPQC College, Waterloo, N2T 2K9 Ontario, Canada
- 25 Istituto Officina dei Materiali IOM-CNR and Dipartimento di Scienze Chimiche e Farmaceutiche, Università degli Studi di Trieste, I-34121 Trieste, Italy
- 26 Department of Physical and Inorganic Chemistry, Universitat Rovira i Virgili, Tarragona, Spain
- 27 ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain
- 28 Chalmers University of Technology, Department of Chemistry and Chemical Engineering, 41296 Gothenburg, Sweden
- 29 Department of Chemistry and Chemical Biology, Department of Physics, and Department of Chemical Engineering, Northeastern University, Boston, Massachusetts 02115, United States
- 30 Department of Chemistry, Fudan University, Shanghai 200433, China
- 31 Institut de Chimie Radicalaire (UMR-7273), Aix-Marseille Univ, CNRS, ICR, 13013 Marseille, France
- 32 Department of Chemistry, Kyungpook National University, Daegu 702-701, South Korea
- 33 Institute of Theoretical Chemistry, Faculty of Chemistry, University of Vienna, Währinger Straße 17, A-1090 Vienna, Austria
- 34 State Key Laboratory of Molecular Reaction Dynamics and Center for Theoretical Computational Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, People's Republic of China
- 35 SSRL, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA
- 36 Department of Chemistry, University of Washington, Seattle, Washington 98195, United States
- 37 Institut für Physik, Universität Rostock, Albert-Einstein-Str. 23-24, 18059 Rostock, Germany
- 38 Institute of Theoretical Physics and Astrophysics, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, ul Wita Stwosza 57, 80-952, Gdańsk, Poland
- 39 Division of Theoretical Chemistry, Chemical Centre, Lund University, P. O. Box 124, SE-22100, Lund, Sweden
- 40 Molecular Sciences Software Institute, Blacksburg, Virginia 24061, United States
- 41 Department of Chemistry, University of Helsinki, P.O. Box 55, FI-00014 University of Helsinki, Finland
- 42 Condensed Matter Theory Group, Institut Néel, CNRS UPR 2940, 38042 Grenoble, France
- 43 Theory Group, Institut Laue Langevin, 38042 Grenoble, France
- 44 Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX 79409-1061, USA
- 45 Department of Chemistry – Ångström Laboratory, Uppsala University, SE-75120 Uppsala, Sweden
- 46 Nantes Université, CNRS, CEISAM, UMR 6230, F-44000 Nantes, France
- 47 Department of Chemistry, National University of Singapore, 3 Science Drive 3, 117543 Singapore
- 48 Current affiliation: School of Physics, Trinity College Dublin, The University of Dublin, College Green, Dublin 2, Ireland.
- 49 Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

- 50 *Department of Chemistry, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8602, Japan*
- 51 *Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan*
- 52 *Department of Chemistry, Loughborough University, Loughborough, LE11 3TU, UK*
- 53 *Quantum Materials and Software LTD, 128 City Road, London, EC1V 2NX, United Kingdom*
- 54 *Scientific Computing Group, Institut Laue Langevin, 38042 Grenoble, France*
- 55 *ENSL, CNRS, Laboratoire de Chimie UMR 5182, 46 allée d'Italie, 69364 Lyon, France*
- 56 *Department of Chemistry and Biochemistry, Butler University, Indianapolis, Indiana 46208, United States*
- 57 *Department of Chemistry, University of California, Irvine, California 92697, United States*
- 58 *Laboratory RA-03, RECENT AIR, A. I. Cuza University of Iași, RA-03 Laboratory (RECENT AIR), Iași 700506, Romania*
- 59 *Chemical Sciences and Engineering Division, Argonne National Laboratory, Lemont, IL 60439, USA*
- 60 *National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6373, USA*
- 61 *Department of Chemistry and Biochemistry, University of Alabama, Tuscaloosa, AL 35487-0336, USA*
- 62 *University Library, University of Southern Denmark, DK-5230 Odense M, Denmark*
- 63 *Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218, USA*
- 64 *Uppsala Center for Computational Chemistry (UC<sub>3</sub>), Uppsala University, PO Box 576, SE-751 23 Uppsala. Sweden*
- 65 *Contributed equally to this manuscript.*

E-mail: g.limanni@fkf.mpg.de; Ignacio.Fernandez@kemi.uu.se; roland.lindh@kemi.uu.se

## Contents

<b>S1</b>	<b>Stochastic Configuration Interaction Eigensolvers</b>	<b>S-7</b>
S1.1	GUGA Spin Adaptation in Stochastic-CASSCF . . . . .	S-7
S1.2	Spin-Purification in SD Basis . . . . .	S-8
S1.3	Stochastic-GASSCF . . . . .	S-10
S1.4	(Spin-)State-Averaged MCSCF . . . . .	S-12
<b>S2</b>	<b>The Heat-Bath Configuration Interaction Self-Consistent Field Method</b>	<b>S-15</b>
S2.1	Computational Details . . . . .	S-15
S2.2	Cartesian Coordinates . . . . .	S-16
S2.3	Active Orbitals . . . . .	S-16
S2.4	Results and Discussion . . . . .	S-18
S2.5	HCI-SCF input example . . . . .	S-20



<b>S3</b>	<b>Uncontracted MRCI</b>	<b>S-20</b>
S3.1	RelaxSE dependencies . . . . .	S-21
S3.2	Output files . . . . .	S-21
S3.3	Input example for CAS+DDCI calculation . . . . .	S-21
S3.4	Input example for SAS+S calculation . . . . .	S-22
S3.5	Symmetry . . . . .	S-23
S3.6	Restart . . . . .	S-23
S3.7	Parallelization . . . . .	S-23
S3.8	Memory . . . . .	S-24
S3.9	Properties calculations after a RelaxSE calculation . . . . .	S-24
<b>S4</b>	<b>Frozen Natural Orbitals method applied to RASPT2</b>	<b>S-24</b>
<b>S5</b>	<b>Derivation of Anisotropic Exchange Interaction from Binuclear Ab Initio Calculations</b>	<b>S-25</b>
<b>S6</b>	<b>Relativistic Hyperfine Coupling</b>	<b>S-40</b>
<b>S7</b>	<b>Single and Two-Photon Spectra with the Exact Semi-Classical Operator</b>	<b>S-42</b>
<b>S8</b>	<b>Vibrational–Rotational Levels for Diatomics</b>	<b>S-47</b>
<b>S9</b>	<b>Evaluating Frenkel’s Excitonic Coupling Terms</b>	<b>S-48</b>
S9.1	Azulene . . . . .	S-48
S9.1.1	Monomer geometries and energetics . . . . .	S-48
S9.1.2	Crystal spectrum evaluation . . . . .	S-56
S9.2	Z–M azobenzene derivatives heterodimer . . . . .	S-58
S9.2.1	Monomer geometries and energetics . . . . .	S-58
S9.2.2	Excitonic spectrum evaluation . . . . .	S-65
S9.3	Plotting absorption spectra . . . . .	S-66
S9.4	Example input files . . . . .	S-66
<b>S10</b>	<b>Ionization and Autoionization Processes</b>	<b>S-69</b>
S10.1	ssDCH . . . . .	S-70
S10.2	Auger Decay Spectra . . . . .	S-72
<b>S11</b>	<b>SA-SI-REKS Quantum–classical Trajectories</b>	<b>S-80</b>
S11.1	Capabilities of the Interface . . . . .	S-80
S11.2	A Technical Description of the Interface . . . . .	S-81
S11.3	Description of Rotary Cycle in Molecular Motors . . . . .	S-84
S11.4	Details of MTDP Simulations in Solvent . . . . .	S-84
<b>S12</b>	<b>Two-Step Acceleration of the Analytic Evaluation of the Vibronic Couplings for Multi-Root Systems</b>	<b>S-85</b>
<b>S13</b>	<b>MC-PDFT Nonadiabatic Dynamics</b>	<b>S-99</b>
<b>S14</b>	<b>ANO-R Basis Set</b>	<b>S-99</b>
S14.1	Hartree–Fock energies for atoms with higher ANO-R- <i>N</i> basis sets . . .	S-99

S14.2	Electron affinity of oxygen atom . . . . .	S-100
<b>S15</b>	<b>Use of SCEPIC program for generation of AIMP</b> <b>input templates for OpenMolcas calculations</b>	<b>S-100</b>
<b>S16</b>	<b>Ab Initio Model Potentials</b>	<b>S-101</b>
S16.1	Quick start . . . . .	S-101
S16.2	Drivers for embedding script . . . . .	S-102
S16.3	Driver for embedding . . . . .	S-103
S16.4	Driver script for SCEPIC . . . . .	S-103
S16.5	Generation of input for OpenMolcas calculation with AIMP	S-104
S16.6	Diffusion of oxygen ad-atom on MgO(001) . . . . .	S-104
<b>References</b>		<b>S-108</b>

# S1 Stochastic Configuration Interaction Eigensolvers

## S1.1 GUGA Spin Adaptation in Stochastic-CASSCF

The OpenMolcas-NECI workflow for GUGA-FCIQMC is displayed in fig. S1. Sample inputs are given in listings S1 and S2.

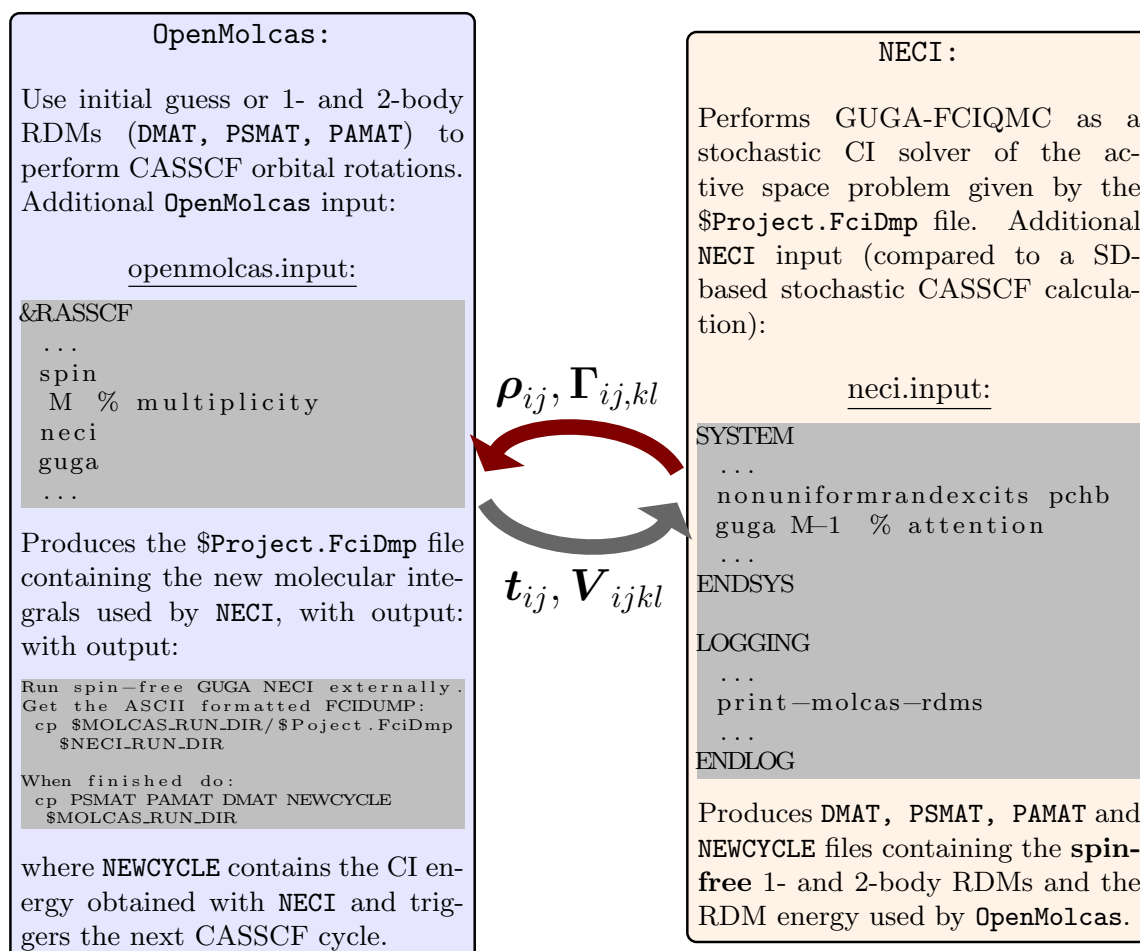


Figure S1: CASSCF/OpenMolcas – GUGA-FCIQMC/NECI Workflow.

Listing S1: Fe<sub>2</sub>S<sub>2</sub> singlet CASSCF OpenMolcas input file. The geometry\_Fe2S2.xyz and fe2s2\_22in26\_ROHF.orbitals files can be found in the SI of ref. S1.

```

1 &GATEWAY
2   coord = geometry_Fe2S2.xyz
3   basis = FE.ANO-RCC-VDZ, ANO-RCC-MB
4   group = nosym
5
6 &SEWARD
7   medium
8
9 &RASSCF
10  file = $CurrDir/fe2s2_22in26_ROHF.orbitals
11  Spin = 1
12  nActEl = 10

```

```

13 Inactive = 88
14 Ras2 = 10
15 guga
16 NECI
17     totalwalkers = 10000
18     RDMLinspace = 200 3 50
19     semistochastic = 1000
20     definedet = 1 3 5 6 9 12 14 16 18 20

```

Listing S2: Fe<sub>2</sub>S<sub>2</sub> singlet CASSCF NECI input file.

```

1 Title Iron-sulfur singlet CASSCF
2
3 System read
4     electrons 10
5     nonuniformrandexcits mol_guga_weighted
6     nobrillouintheorem
7     guga 0
8 endsys
9
10 calc
11     methods
12         method vertex fcimc
13     endmethods
14
15     definedet 1 3 5 7 9 12 14 16 18 20
16     totalwalkers 100000
17     semi-stochastic 5000
18     pops-core 10000
19     truncinitiator
20     addtoinitiator 3
21     allrealcoeff
22     tau-values \
23         start user-defined 0.01 \
24         max 0.02
25     tau-search \
26         algorithm histogramming
27     startsinglepart 100
28     rdmsamplingiters 10000
29 endcalc
30
31 logging
32     hdf5-pops
33     print-molcas-rdms
34     calcrdmonfly 3 5000 100
35 endlog
36 end

```

## S1.2 Spin-Purification in SD Basis

Sample inputs for an open-shell singlet calculation are given in listings S3 and S4

Listing S3: Preparation of the oxygen <sup>1</sup>Δ state in OpenMolcas for NECI.

```

1 &GATEWAY
2     coord = O2.xyz
3     basis = ANO-RCC-VDZP

```

```

4      group = C1
5      RICD
6
7  &SEWARD
8
9  &RASSCF
10     file = $CurrDir/02.InpOrb
11     spin = 1
12     /* (16, 28) is Full CI in VDZP */
13     nActEl = 16
14     RAS2 = 28
15     CIOnly
16     NECI
17     totalwalkers = 10000000
18     RDMLinspace = 200000 20 1000

```

Listing S4: (16, 28) Full CI on the oxygen  $^1\Delta$  state in NECI.

```

1  title Oxygen singlet delta state
2
3  system read
4     electrons 16
5     nonuniformrandexcits pchb
6     nobrillouintheorem
7     spin-restrict 0
8     FCIDUMP-name 02.FciDmp
9     freeformat
10
11     sd-spin-purification 0.1185
12 endsys
13
14 calc
15     time 1380
16
17     definedet \
18         1-2 3-4 \
19         5-6 7-8 \
20         9-10 11-12 13-14 \
21         15 18
22
23     totalwalkers 5.e7
24     readpops
25     walkcontgrow
26
27     semi-stochastic 5000
28     pops-core 10000
29
30     methods
31         method vertex fcimc
32     endmethods
33
34     diagshift .00
35     shiftdamp .02
36     stepsshift 10
37     proje-changeref 1.5
38     truncinitiator
39     addtoinitiator 3
40     allrealcoeff

```

```

41  realspawncutoff .30
42  jump-shift
43  tau-values \
44      start user-defined 0.01
45  tau-search \
46      algorithm conventional \
47      stop-condition no-change 2000 \
48      maxwalkerbloom 1
49  memoryfacspawn 100.00
50  memoryfacpart 50.00
51  startsinglepart 10
52  endcalc
53
54  logging
55      highlypopwrite 50
56      hdf5-pops
57      print-spin-resolved-RDMs
58      printonerdm
59      RDMlinspace 200000 20 1000
60  endlog
61  end

```

### S1.3 Stochastic-GASSCF

Sample input files for a Stochastic-GASSCF calculation are provided in listings S5 and S6

Listing S5: Preparation of a GASSCF calculation on a stack of five benzene molecules with single interspace excitations in OpenMolcas for NECI.

```

1  &GATEWAY
2      RICD
3      coord = $Project.xyz
4      basis = ANO-RCC-VDZP
5      group = C1
6
7  &SEWARD
8
9
10 &RASSCF
11     THRS = 1.0e-4 1.0e-1 5.0e-4
12     FileOrb = $CurrDir/$Project.InpOrb
13     nActEl = 30
14     Symmetry = 1
15     Spin = 1
16     GASSCF
17         5
18         6
19         5 7
20         6
21         11 13
22         6
23         17 19
24         6
25         23 25
26         6
27         30 30
28

```

```

29 * Switch on NECI (external by default)
30 NECI
31   totalwalkers = 1000000
32   RDMLinspace = 30000 20 500
33   semistochastic = 20000
34   definedet = 1 2 3 4 5 6      13 14 15 16 17 18      25 26 27 28 29 30      37 38
      ↪ 39 40 41 42      49 50 51 52 53 54

```

Listing S6: GASSCF on a benzene stack with single interspace excitations.

```

1  Title GAS on a stack of 5 benzene
2
3  System read
4    electrons 30
5    nobrillouintheorem
6    nonuniformrandexcits GAS-CI PCHB
7    GAS-spec cumulative 5
8      6 5 7
9      6 11 13
10     6 17 19
11     6 23 25
12     6 30 30
13     6*1 6*2 6*3 6*4 6*5
14   FCIDUMP-name benzene_3.0_Ang.FciDmp
15 endsys
16
17 calc
18   definedet 1-6 13-18 25-30 37-42 49-54
19
20   totalwalkers 1e6
21
22   semi-stochastic 20000
23   pops-core 10000
24
25   methods
26     method vertex fcimc
27   endmethods
28
29   shiftdamp .02
30   stepsshift 10
31
32   tau-values \
33     start user-defined 5e-3
34
35   tau-search \
36     algorithm conventional \
37     stop-condition no-change 2000 \
38     maxwalkerbloom 1
39
40   nmcyc 10000
41
42   proje-changeref 1.20
43   truncinitiator
44   addtoinitiator 3
45   allrealcoeff
46   realspawncutoff .30
47   memoryfacspawn 10.00
48   memoryfacpart 5.00

```



```

49     startsinglepart 10
50
51 endcalc
52
53 logging
54     highlypopwrite 50
55     hdf5-pops
56     print-spin-resolved-RDMs
57     printonerdm
58     RDMLinspace 30000 20 500
59 endlog
60 end

```

## S1.4 (Spin-)State-Averaged MCSCF

A sample input for a MCSCF calculation over different spin multiplicities is given in listings S7 and S8.

Listing S7: Spin-Averaged CASSCF over doublet, quartet, sextet and octet multiplicities.

```

1 &GATEWAY
2     coord = mn2o4.xyz
3     basis = ANO-RCC-VDZP
4     group = nosym
5     RICD
6
7 &SEWARD
8
9 &RASSCF
10    FileOrb = $Project.InpOrb
11    CIRROOT = 4 4 1
12        * average over four roots of arbitrary multiplicity,
13        * alternatively, to control the root weights w_i, use
14        * CIRROOT = 4 4
15        *           1   2   3   4
16        *           w_1 w_2 w_3 w_4
17    spin = 2
18        * multiplicity specified here will be ignored
19    nactel = 19 0 0
20    inactive = 118
21    THRS = 1.0e-07, 1.0e-04, 1.0e-04
22    ras2 = 16
23    deleted = 0
24    Symmetry = 1
25    ITERation = 200 100
26    prwf = 0.01
27    orbllisting = all
28    NECI
29        GUGA
30        totalwalkers = 20000
31        RDMLinspace = 20000 10 100

```

Listing S8: NECI input for the doublet spin state.

```

1 # With the FCIDUMP generated within Molcas, separate NECI dynamics need to be
2 # run for each spin multiplicity. If one spin sector
3 # contains multiple states, these can be computed within

```

```

4 # one dynamic through the "system replica" keyword.
5
6 title Mn204 spin-averaging doublet
7 System read
8   electrons          19
9   guga                1
10  nonuniformrandexcits guga-pchb
11  system-replicas    2      # twice the number of doublet states
12  nobrillouintheorem
13  freeformat
14 Endsys
15
16 Calc
17  methods
18    method vertex fcimc
19  endmethods
20  totalWalkers      2.5e6
21
22  definedet         1 2 3 4 5 6 7 8 9 10 11 12      13 15 17 21      24 26 28
23  multiple-initial-states
24    1 2 3 4 5 6 7 8 9 10 11 12      13 15 17 21 24 26 28
25    1 2 3 4 5 6 7 8 9 10 11 12      13 15 17 21 24 26 28
26
27  global-core-space
28  orthogonalise-replicas
29  replica-single-det-start
30  semi-stochastic 10000
31  pops-core        100000
32  rdmsamplingiters 40000
33  time 12000
34  nmcyc 1000000
35
36  tau-values \
37    start user-defined 0.01
38  tau-search \
39    algorithm conventional \
40    stop-condition no-change 2000 \
41    maxwalkerbloom 1
42
43  truncate-spawns 2.0
44  diagshift 0.00
45  shiftdamp 0.3
46  stepsshift 10
47  proje-changeref 1.2
48  startsinglepart 100
49  memoryfacspawn 100.0
50  memoryfacpart 5.0
51  truncinitiator
52  addtoinitiator 3
53  allrealcoeff
54  realspawncutoff 0.30
55 Endcalc
56
57 Logging
58  highlypopwrite      2000
59  hdf5-pops
60  calcrdmonfly        3 15000 100
61  print-molcas-rdms

```

```
62 Endlog
63 End
```

A sample input for a Deterministic Spin-Averaged CASSCF over one doublet and a quartet state is given in listing S9

Listing S9: Deterministic Spin-Averaged CASSCF over one doublet and a quartet state.

```
1
2 * Three separate (!) input files are required for the deterministic CASSCF over
3 * two multiplicities. Below three sample &RASSCF blocks are provided:
4
5 * 1. CAS-CI for the doublet:
6 &RASSCF
7   CIONly
8   FileOrb = $Project.InpOrb
9   CIRoot = 1 1 1
10  spin = 2
11  nactel = 19 0 0
12  ras2 = 16
13  inactive = 118
14  WRMA
15      * this keyword writes the PSMAT, PAMAT, DMAT, DSPN arrays to $WORKDIR
16      * which represent the symmetric- and anti-symmetric part of the
17      * 2RDM, the 1RDM and the 1 Spin-Density Matrix
18
19 * 2. CAS-CI for the quartet
20 &RASSCF
21   CIONly
22   FileOrb = $Project.InpOrb
23   CIRoot = 1 1 1
24   spin = 4
25   nactel = 19 0 0
26   ras2 = 16
27   inactive = 118
28   WRMA
29      * Another set of RDMs for the quartet
30
31 * 3. CASSCF which averages the RDMs
32 &RASSCF
33   FileOrb = $Project.InpOrb
34   CIRoot = 2 2 1
35      * average over two roots of arbitrary multiplicity,
36   spin = 2
37      * multiplicity specified here will be ignored
38   nactel = 19 0 0
39   inactive = 118
40   ras2 = 16
41   NECI
42     GUGA
43     totalwalkers = 20000
44     RDMLinspace = 20000 10 100
45     * Specification of the entire NECI block as shown here is
46     * mandatory. The program will ask you to provide the RDMs generated
47     * above in consecutive numbering and a NEWCYCLE file which contains
48     * the CAS-CI energies. Any renaming to PSMAT.1, PSMAT.2, ... has
49     * to be manually performed by the user. Please consult the
50     * GUGA-NECI interface for further guidance.
51     * For each iteration of the SCF, the steps (1) and (2) have to be
```

## S2 The Heat-Bath Configuration Interaction Self-Consistent Field Method

Application of the HCI-SCF Method:  $[(C_5H_5)Fe(C_5H_4)]_2^+$ .

### S2.1 Computational Details

- Two active spaces were employed, CAS(27,28) and CAS(35,36). The active natural orbitals are shown in figs. S2 and S3.
- Cartesian coordinates were obtained at the B3LYP/def2-TZVP level of theory (listing S10). All B3LYP calculations were done with the `Turbomole` package.<sup>S2</sup>
- The  $C_{2h}$  point group symmetry was used in all calculations.
- cc-pVTZ basis set for Fe, and cc-pVDZ basis set for C and H were employed.
- The Cholesky decomposition technique for the two-electron integrals was used, with a threshold of  $10^{-6} E_h$ .
- DMRG-SCF calculations were done with the `OpenMolcas-CheMPS2` interface. Fiedler orbital ordering, perturbative noise (noise prefactor of 0.05), and residual norm threshold  $10^{-4}$  for the Davidson algorithm were employed. Four values of  $m$  were chosen, namely 500, 1000, 2000, and 4000.
- HCI-SCF calculations were done with the `OpenMolcas-Dice` interface. We performed three sets of calculations with  $\varepsilon_{\text{var}} = 10^{-4} E_h$ ,  $5 \cdot 10^{-5} E_h$ ,  $10^{-4} E_h$  and  $\varepsilon_{\text{per}} = 10^{-5} E_h$ ,  $5 \cdot 10^{-6} E_h$ ,  $10^{-6} E_h$ , respectively. The last calculation is semi-stochastic, in which the perturbative component is calculated deterministically at  $\varepsilon = 10^{-5} E_h$  and stochastically at  $\varepsilon = 10^{-6} E_h$ . Since `Dice` does not calculate the 2-RDM with semi-stochastic HCI, the active orbitals were not optimized (activated by the keyword `CIONly`). The active orbitals in this last calculation were taken from the HCI-SCF calculations with  $\varepsilon_{\text{var}} = 10^{-4} E_h$  and  $\varepsilon_{\text{per}} = 10^{-5} E_h$ .
- All calculations were done with an Intel(R) Xeon(R) CPU E5-2650 @ 2.00 GHz (8 cores, 16 threads) and 62 GB of memory. `CheMPS2` was built with a shared-memory OMP parallelization, whereas `Dice` was built with MPI parallelization. The number of `OpenMolcas` processes is 4 (`MOLCAS_NPROCS=4`), the number of `CheMPS2` (or `Dice`) processes is 12 (`OMP_NUM_THREADS=12` and `MOLCAS_DICE=12`)
- The electronic coupling  $V_{\text{ET}}$  was calculated using the generalized Mulliken-Hush method:<sup>S3</sup>

$$V_{\text{ET}} = \frac{(E_2 - E_1) |\boldsymbol{\mu}_{12}|}{\sqrt{(\boldsymbol{\mu}_{11} - \boldsymbol{\mu}_{22})^2 + 4 |\boldsymbol{\mu}_{12}|^2}} \quad (\text{S1})$$

with  $E_2$  the energy of the excited state  ${}^2B_u$ ,  $E_1$  the energy of the ground state  ${}^2A_g$ ,  $\boldsymbol{\mu}_{22}$  the electric dipole moment of  ${}^2B_u$ ,  $\boldsymbol{\mu}_{11}$  the electric dipole moment of  ${}^2A_g$ , and

$\mu_{12}$  the transition dipole moment. Since  $[(C_5H_5)Fe(C_5H_4)]_2^+$  has an inversion center, eq. (S1) becomes:

$$V_{ET} = \frac{E_2 - E_1}{2} \quad (S2)$$

## S2.2 Cartesian Coordinates

Listing S10: B3LYP/def2-TZVP structure (in Å) of  $[(C_5H_5)Fe(C_5H_4)]_2^+$ .

1	40			
2				
3	C	1.9982613	2.1397439	0.7101555
4	C	1.0606557	1.1702929	1.1519365
5	C	0.5036756	0.5158557	0.0000000
6	C	1.0606557	1.1702929	-1.1519365
7	C	1.9982613	2.1397439	-0.7101555
8	H	2.5875873	2.7816599	1.3448739
9	H	0.8282172	0.9447440	2.1799754
10	H	0.8282172	0.9447440	-2.1799754
11	H	2.5875873	2.7816599	-1.3448739
12	Fe	0.0992538	2.6229777	0.0000000
13	C	-1.9749164	3.0878516	0.0000000
14	C	-1.3541382	3.6402869	-1.1501836
15	C	-0.3477934	4.5468736	-0.7107298
16	C	-0.3477934	4.5468736	0.7107298
17	C	-1.3541382	3.6402869	1.1501836
18	H	-2.7591408	2.3470637	0.0000000
19	H	-1.5980451	3.4116585	-2.1751106
20	H	0.3040348	5.1258255	-1.3449579
21	H	0.3040348	5.1258255	1.3449579
22	H	-1.5980451	3.4116585	2.1751106
23	H	1.5980451	-3.4116585	-2.1751106
24	H	-0.8282172	-0.9447440	-2.1799754
25	H	-2.5875873	-2.7816599	-1.3448739
26	C	1.3541382	-3.6402869	-1.1501836
27	C	-1.0606557	-1.1702929	-1.1519365
28	C	-1.9982613	-2.1397439	-0.7101555
29	H	2.7591408	-2.3470637	0.0000000
30	C	1.9749164	-3.0878516	0.0000000
31	C	0.3477934	-4.5468736	-0.7107298
32	H	-0.3040348	-5.1258255	-1.3449579
33	Fe	-0.0992538	-2.6229777	0.0000000
34	C	-0.5036756	-0.5158557	0.0000000
35	C	-1.9982613	-2.1397439	0.7101555
36	C	1.3541382	-3.6402869	1.1501836
37	C	0.3477934	-4.5468736	0.7107298
38	H	-2.5875873	-2.7816599	1.3448739
39	C	-1.0606557	-1.1702929	1.1519365
40	H	1.5980451	-3.4116585	2.1751106
41	H	-0.3040348	-5.1258255	1.3449579
42	H	-0.8282172	-0.9447440	2.1799754

## S2.3 Active Orbitals

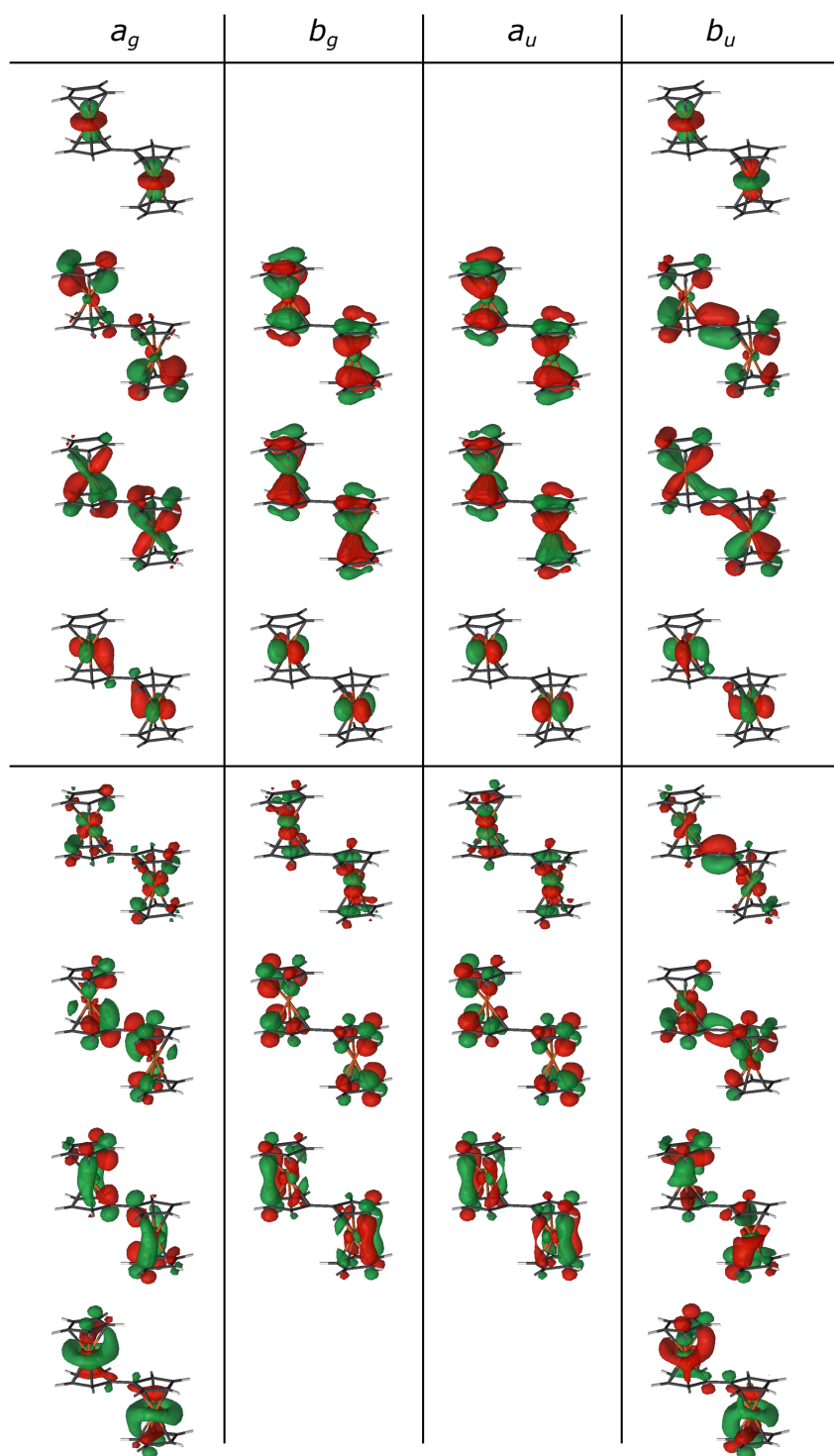


Figure S2: CAS(27,28) active natural orbitals of  $[(C_5H_5)Fe(C_5H_4)]_2^+$ .

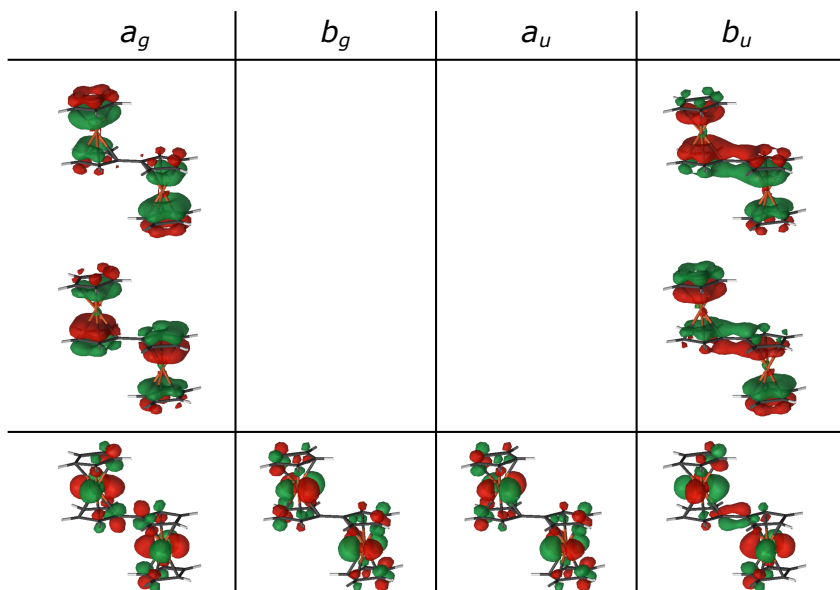


Figure S3: Extra active orbitals in the CAS(35,36) calculations of  $[(C_5H_5)Fe(C_5H_4)]_2^+$ .

## S2.4 Results and Discussion

In tables S1 and S2, we present the total energies (in  $E_h$ ) of the ground state  $^2A_g$  and the excited state  $^2B_u$  of  $[(C_5H_5)Fe(C_5H_4)]_2^+$ , calculated with DMRG-SCF and HCI-SCF. It can be seen that the total energies are far from converged in all cases. In contrast, the electronic coupling  $V_{ET}$  converges relatively fast with respect to the number of renormalized states  $m$  and the thresholds  $\varepsilon$ . For the smaller active space CAS(27,28), HCI-SCF, DMRG-SCF [ $m = 4000$ ], and semi-stochastic HCI predict a  $V_{ET}$  value of  $\sim 4.4 \text{ kcal mol}^{-1}$ . For the larger active space CAS(35,36), even at an  $m$ -value of 6000, the electronic coupling has not yet converged. HCI-SCF and semi-stochastic HCI, on the other hand, give very similar results ( $4.4 \text{ kcal mol}^{-1}$ ). Since there is still a gap of  $1.2 \text{ kcal mol}^{-1}$  between the best (S)HCI and DMRG results, the converged relative energy is expected to be somewhere between  $4.5 \text{ kcal mol}^{-1}$  and  $5.7 \text{ kcal mol}^{-1}$ .



Table S1: Total energies (in  $E_h$ ) of the ground state  ${}^2A_g$  and the excited state  ${}^2B_u$  of  $[(C_5H_5)Fe(C_5H_4)]_2^+$ , calculated with DMRG-SCF(27,28) and HCI-SCF(27,28).

	$\varepsilon_{\text{var}}$	$\varepsilon_{\text{per}}$	$E({}^2A_g)$	$E({}^2B_u)$	$V_{\text{ET}}^a$
	$(E_h)$		$(E_h, -3292)$		$(\text{kcal mol}^{-1})$
HCI-SCF	$10^{-4}$	$10^{-5}$	-0.799 611	-0.784 802	4.6
HCI-SCF	$5 \cdot 10^{-5}$	$5 \cdot 10^{-6}$	-0.802 740	-0.789 133	4.3
HCI <sup>b</sup>	$10^{-4}$	$10^{-6}$	-0.808 930	-0.794 811	4.4
DMRG-SCF[ $m = 500$ ]			-0.798 354	-0.781 977	5.1
DMRG-SCF[ $m = 1000$ ]			-0.803 205	-0.792 376	3.4
DMRG-SCF[ $m = 2000$ ]			-0.807 224	-0.795 179	3.8
DMRG-SCF[ $m = 4000$ ]			-0.810 455	-0.796 727	4.3

<sup>a</sup> The electronic coupling  $V_{\text{ET}}$  is half of the energy gap between  ${}^2A_g$  and  ${}^2B_u$ .

<sup>b</sup> Semi-stochastic HCI calculation, the perturbative component is calculated deterministically at  $\varepsilon = 10^{-5} E_h$  and stochastically at  $\varepsilon = 10^{-6} E_h$ . The active space is taken from the cheapest HCI-SCF calculation.

Table S2: Total energies (in  $E_h$ ) of the ground state  ${}^2A_g$  and the excited state  ${}^2B_u$  of  $[(C_5H_5)Fe(C_5H_4)]_2^+$ , calculated with DMRG-SCF(35,36) and HCI-SCF(35,36).

	$\varepsilon_{\text{var}}$	$\varepsilon_{\text{per}}$	$E({}^2A_g)$	$E({}^2B_u)$	$V_{\text{ET}}^a$
	$(E_h)$		$(E_h, -3292)$		$(\text{kcal mol}^{-1})$
HCI-SCF <sup>b</sup>	$10^{-4}$	$10^{-5}$	-0.888 687	-0.874 637	4.4
HCI <sup>c</sup>	$10^{-4}$	$10^{-6}$	-0.908 208	-0.893 885	4.5
DMRG-SCF[ $m = 500$ ]			-0.884 200	-0.870 386	4.3
DMRG-SCF[ $m = 1000$ ]			-0.906 087	-0.879 654	8.3
DMRG-SCF[ $m = 2000$ ]			-0.912 357	-0.890 646	6.8
DMRG-CI[ $m = 4000$ ] <sup>d</sup>			-0.919 855	-0.900 708	6.0
DMRG-CI[ $m = 6000$ ] <sup>d</sup>			-0.922 990	-0.904 974	5.7

<sup>a</sup> The electronic coupling  $V_{\text{ET}}$  is half of the energy gap between  ${}^2A_g$  and  ${}^2B_u$ .

<sup>b</sup> HCI-SCF calculations with  $\varepsilon_{\text{var}} = 5 \cdot 10^{-6} E_h$  require more than 62 GB of memory.

<sup>c</sup> Semi-stochastic HCI calculation, the perturbative component is calculated deterministically at  $\varepsilon = 10^{-5} E_h$  and stochastically at  $\varepsilon = 10^{-6} E_h$ . The active space is taken from the cheapest HCI-SCF calculations.

<sup>d</sup> DMRG-CI calculations only, the active space is taken from the DMRG-SCF[ $m = 2000$ ] calculations.

## S2.5 HCI-SCF input example

In listing S11, we present an input example to perform an HCI-CAS calculation. This input is to calculate the  ${}^2A_g$  state of  $[(C_5H_5)Fe(C_5H_4)]_2^+$ . The keyword *DICE* must be used to activate the *OpenMolcas-Dice* interface. *DIOOccupations* is used to specify a starting (or reference) determinant. The keyword requires first the number of determinants, followed by determinants. Each determinant is inserted as a string of aliases of occupations/couplings of the active (RAS2) orbitals with the aliases *2* = full, *u* = up, *d* = down, *o* = empty. Dice can run in parallel with the number of Dice processes set by *MOLCAS\_DICE*.

Listing S11: HCI-CAS(27,28) calculation of the  ${}^2A_g$  state of  $[(C_5H_5)Fe(C_5H_4)]_2^+$ .

```
1 >> export MOLCAS_DICE=12
2
3 &RASSCF
4 SPIN      = 1
5 SYMMetry = 2
6 NACTel    = 27 0 0
7 INACTive  = 14 8 8 13
8 RAS2      = 8 6 6 8
9 THRS      = 1.0e-06 1.0e-03 1.0e-03
10 LUMOrb
11 DICE
12 EPSilon  = 1.0e-4 1.0e-5
13 DIOOccupations
14 1
15 2 2 2 u 0 0 0 0 2 2 2 0 0 0 2 2 2 0 0 0 2 2 2 2 0 0 0 0
```

In listing S12, we show the same calculation but with semi-stochastic HCI.

Listing S12: Semi-stochastic HCI(27,28) calculation of the  ${}^2A_g$  state of  $[(C_5H_5)Fe(C_5H_4)]_2^+$ .

```
1 >> export MOLCAS_DICE=12
2
3 &RASSCF
4 SPIN      = 1
5 SYMMetry = 2
6 NACTel    = 27 0 0
7 INACTive  = 14 8 8 13
8 RAS2      = 8 6 6 8
9 THRS      = 1.0e-06 1.0e-03 1.0e-03
10 LUMOrb
11 DICE
12 EPSilon  = 1.0e-4 1.0e-6
13 DIOOccupations
14 1
15 2 2 2 u 0 0 0 0 2 2 2 0 0 0 2 2 2 0 0 0 2 2 2 2 0 0 0 0
16 STOChastic
17 CIONly
```

## S3 Uncontracted MRCI

Use of the *RelaxSE* code for MRCI calculations.

This section describes typical inputs for CAS+DDCI and SAS+S types of calculations. It also details extra information on parallelization, symmetry, restart, and the property associated code.

### S3.1 RelaxSE dependencies

The RelaxSE code requires

- the TraOne and TraInt files issued from the MOTRA module in OpenMolcas, computed in the conventional way,
- an INPUT file providing the input namelists,
- a prefix.ref0 input file giving the configurations selected within the CAS when required.

### S3.2 Output files

The RelaxSE code provides the following files

- the output file as prefix.out,
- the wave functions at the last Davidson iteration (for restart or properties calculations) in prefix.restart,

and if required by the presence of *propinp* namelist in the INPUT file

- the prefix.bdet file containing the list of determinants,
- the prefix.sass file containing extra information.

### S3.3 Input example for CAS+DDCI calculation

This calculation has been published in ref. S4.

Listing S13: INPUT file.

```
1  &sassinp
2    prefix="clusters1",
3    method = "SAS+S"
4    restart = .false.
5  &end
6
7  &infomolcasinp
8    nirrep=1,
9    ntot=235,
10   iIrTb(1,1) = 1,
11   iChTb(1,1) = 1
12 &end
13
14 &orbinp
15   norb_gel = 2,
16   norb_occ = 52,
17   norb_ligo = 4,
18   norb_act = 4,
19   norb_ligv = 0,
```

```

20     norb_virt = 175,
21     norb_del  = 11,
22 &end
23
24 &vecinp
25     stot=3,
26     sz=2,
27     vec_irrep=1,
28     nvec=2,
29     nref0=1,
30     nelact = 4
31 &end
32
33 &davidinp
34     SizeheffDavidson = 10,
35     tol_conv = 1.d-8,
36     ITER0 = 1
37 &end
38
39 &propinp
40     Nprop=1
41 &end

```

Listing S14: clusters1.sass.ref0 file.

```

1 59u 60u 61u 62d

```

### S3.4 Input example for SAS+S calculation

This calculation has been published in ref. S5.

Listing S15: INPUT file.

```

1 &sassinp
2     prefix="Joo2",
3     method = "SAS+S"
4     restart = .false.
5 &end
6
7 &infomolcasinp
8     nirrep=1,
9     ntot=248,
10     iIrTb(1,1) = 1,
11     iChTb(1,1) = 1
12 &end
13
14
15 &orbinp
16     norb_gel = 0,
17     norb_occ = 50,
18     norb_ligo = 6,
19     norb_act = 6,
20     norb_ligv = 0,
21     norb_virt = 186,
22     norb_del = 0,
23 &end
24

```

```

25 &vecinp
26   stot=5,
27   sz=4,
28   vec_irrep=1,
29   nvec=2,
30   nref0=1,
31   nelact = 6
32 &end
33
34 &davidinp
35 SizeheffDavidson = 10,
36 tol_conv = 1.d-8,
37 ITER0 = 1
38 &end
39
40 &propinp
41   Nprop=1
42 &end

```

Listing S16: Joo2.sass.ref0 file.

```

1 57u 58u 59u 60u 61u 62d

```

### S3.5 Symmetry

In the present version of the the `RelaxSE` code, the symmetry is not implemented. As a consequence the symmetry information in the `infomolcasinp` namelist should be left as in the previous examples.

### S3.6 Restart

If the calculation is not finished while stopped for any reason, or if one desire more accuracy, a restart from the last Davidson's iteration can be done. The `prefix.restart` file will then be required in addition to the `INPUT` and `prefix.ref0` files. To specify to the code to restart from the `prefix.restart` file one should add the following keywords:

- in the `sassinp` namelist: `restart=.true.` ,
- in the `davidinp` namelist: `iter0=number_of_the_last_iteration` in the previous run, the `tol_conv` keyword can also be specified to increase the accuracy of the convergence for the Davidson's diagonalization procedure.

### S3.7 Parallelization

The `RelaxSE` code is implemented in a MPI+OpenMP scheme. The number of MPI processes should remain small (4–6 is an optimum) while the number of OpenMP threads per MPI processes has been used up to 40 with a good scaling (see ref. S6) for more precision. It is also recommended to adjust the optional keyword `SIZEBATCH` (default 40) to a multiple of the number of OpenMP threads, as close as possible to a divisor of the number of virtual+ligv orbitals.

## S3.8 Memory

The RelaxSE use of memory is often a problem for large calculations. For up to  $250 \cdot 10^6$  determinants, it should work on a computer with 200 GB of memory per node. When reaching  $10^9$  determinants one should have more than 1 TB of memory per node available.

A low memory version of the code is under progress.

## S3.9 Properties calculations after a RelaxSE calculation

A property code is provided in the RelaxSE realise. To be used it requires the same INPUT as RelaxSE, the `prefix.restart`, `prefix.bdet` and `prefix.sass` file. The last ones are generated by RelaxSE only if the *propinp* namelist is present in the INPUT file. If forgotten in the first run a restart calculation can be done to generate them.

Presently only few properties are coded,

- the projection of the computed states wave functions into the CAS space
- the calculation of the one-electron density matrices,
- the printing of the computed states wave functions in human readable format (to be used with care).

Other properties implementations are in progress.

## S4 Frozen Natural Orbitals method applied to RASPT2

An input example for one of the FNO-RASPT2 calculations reported on the PSB-11 system ( $\zeta = 99\%$ ,  $\sigma = 0.1 E_h$ ) is given in listing S17

Listing S17: OpenMolcas input for a FNO-RASPT2 calculation.

```
1 &gateway
2   Coord=$HomeDir/coord.xyz
3   Basis=ANO-L-VDZP
4   Group=nosymm
5   RICD
6
7 &seward
8
9 >>>COPY $HomeDir/cas.RasOrb INPORB
10
11 &rasscf
12   spin
13   1
14   symm
15   1
16   nactel
17   12 3 3
18   ras1
19   6
20   ras2
21   0
22   ras3
```

```

23      6
24      inactive
25      76
26      ciroot
27      3 3 1
28      lumorb
29
30
31      &caspt2
32      imag=0.2
33      ipea=0.0
34      maxiter=200
35      multistate=3 1 2 3
36      nomult
37      FNOCaspt2=-0.01
38      RegFN0parameter=0.1

```

## S5 Derivation of Anisotropic Exchange Interaction from Binuclear Ab Initio Calculations

Table S3 collects all parameters of the anisotropic magnetic exchange extracted from calculation of the binuclear system.

Table S3: Parameters of the anisotropic magnetic exchange (in  $\text{cm}^{-1}$ ) extracted from calculation of the DyMn binuclear system. The active space of the CASSCF method included  $4f^9$  and  $5d^5$  shells of the  $\text{Dy}^{\text{III}}$  and  $\text{Mn}^{\text{II}}$ , respectively, amounting to 14 electrons in 12 orbitals. All roots arising from the coupling of the ground  $^6\text{H}$  term of  $\text{Dy}^{\text{III}}$  and ground spin  $S_B = 5/2$  of  $\text{Mn}^{2+}$  were explicitly optimized and mixed by spin-orbit interaction in RASSI. The ANO-RCC-VTZP basis set was used for closer atoms, while smaller VDZP contractions were used for distant atoms. This is the full version of table 1

$k_A$	$q_A$	$k_B$	$q_B$	Real Part	Imaginary Part
1	-1	1	-1	$-7.3853 \cdot 10^{-5}$	$9.0536 \cdot 10^{-5}$
1	-1	1	0	$-3.4434 \cdot 10^{-5}$	$2.7705 \cdot 10^{-6}$
1	-1	1	1	$-4.7023 \cdot 10^{-4}$	$1.9874 \cdot 10^{-4}$
1	-1	3	-3	$2.2467 \cdot 10^{-8}$	$-1.3965 \cdot 10^{-8}$
1	-1	3	-2	$9.2186 \cdot 10^{-8}$	$1.8182 \cdot 10^{-7}$
1	-1	3	-1	$-3.7601 \cdot 10^{-7}$	$2.8688 \cdot 10^{-7}$
1	-1	3	0	$-7.3376 \cdot 10^{-7}$	$-1.0045 \cdot 10^{-6}$
1	-1	3	1	$2.2679 \cdot 10^{-6}$	$1.9140 \cdot 10^{-6}$
1	-1	3	2	$-4.2336 \cdot 10^{-7}$	$-8.4868 \cdot 10^{-7}$
1	-1	3	3	$1.9361 \cdot 10^{-5}$	$-4.4965 \cdot 10^{-5}$
1	-1	5	-5	$-3.4628 \cdot 10^{-11}$	$4.2841 \cdot 10^{-11}$
1	-1	5	-4	$-1.4420 \cdot 10^{-10}$	$9.1725 \cdot 10^{-11}$
1	-1	5	-3	$2.4222 \cdot 10^{-9}$	$-6.7578 \cdot 10^{-10}$
1	-1	5	-2	$3.2058 \cdot 10^{-9}$	$-1.9943 \cdot 10^{-9}$
1	-1	5	-1	$2.1029 \cdot 10^{-10}$	$1.1379 \cdot 10^{-8}$
1	-1	5	0	$-1.6505 \cdot 10^{-8}$	$8.9115 \cdot 10^{-9}$
1	-1	5	1	$-2.5302 \cdot 10^{-8}$	$-8.2347 \cdot 10^{-8}$



1	-1	5	2	$1.7726 \cdot 10^{-7}$	$-5.9765 \cdot 10^{-9}$
1	-1	5	3	$2.2294 \cdot 10^{-7}$	$1.0088 \cdot 10^{-6}$
1	-1	5	4	$-1.0777 \cdot 10^{-6}$	$-3.3309 \cdot 10^{-7}$
1	-1	5	5	$-2.2523 \cdot 10^{-6}$	$-4.2296 \cdot 10^{-6}$
1	0	1	-1	$2.0764 \cdot 10^{-2}$	$-1.2782 \cdot 10^{-3}$
1	0	1	0	-1.6573	$2.6958 \cdot 10^{-18}$
1	0	1	1	$-2.0764 \cdot 10^{-2}$	$-1.2782 \cdot 10^{-3}$
1	0	3	-3	$7.1690 \cdot 10^{-5}$	$-7.3159 \cdot 10^{-5}$
1	0	3	-2	$7.0613 \cdot 10^{-4}$	$1.7459 \cdot 10^{-5}$
1	0	3	-1	$7.2281 \cdot 10^{-5}$	$3.2494 \cdot 10^{-4}$
1	0	3	0	$-5.8643 \cdot 10^{-3}$	$-4.3972 \cdot 10^{-19}$
1	0	3	1	$-7.2281 \cdot 10^{-5}$	$3.2494 \cdot 10^{-4}$
1	0	3	2	$7.0613 \cdot 10^{-4}$	$-1.7459 \cdot 10^{-5}$
1	0	3	3	$-7.1690 \cdot 10^{-5}$	$-7.3159 \cdot 10^{-5}$
1	0	5	-5	$9.0205 \cdot 10^{-8}$	$-5.9258 \cdot 10^{-9}$
1	0	5	-4	$8.1000 \cdot 10^{-7}$	$5.3169 \cdot 10^{-7}$
1	0	5	-3	$-1.1102 \cdot 10^{-6}$	$7.4078 \cdot 10^{-7}$
1	0	5	-2	$-7.3518 \cdot 10^{-6}$	$-5.8876 \cdot 10^{-7}$
1	0	5	-1	$-3.2085 \cdot 10^{-6}$	$-7.3596 \cdot 10^{-6}$
1	0	5	0	$1.5686 \cdot 10^{-5}$	$-4.7561 \cdot 10^{-21}$
1	0	5	1	$3.2085 \cdot 10^{-6}$	$-7.3596 \cdot 10^{-6}$
1	0	5	2	$-7.3518 \cdot 10^{-6}$	$5.8876 \cdot 10^{-7}$
1	0	5	3	$1.1102 \cdot 10^{-6}$	$7.4078 \cdot 10^{-7}$
1	0	5	4	$8.1000 \cdot 10^{-7}$	$-5.3169 \cdot 10^{-7}$
1	0	5	5	$-9.0205 \cdot 10^{-8}$	$-5.9258 \cdot 10^{-9}$
1	1	1	-1	$-4.7023 \cdot 10^{-4}$	$-1.9874 \cdot 10^{-4}$
1	1	1	0	$3.4434 \cdot 10^{-5}$	$2.7705 \cdot 10^{-6}$
1	1	1	1	$-7.3853 \cdot 10^{-5}$	$-9.0536 \cdot 10^{-5}$
1	1	3	-3	$1.9361 \cdot 10^{-5}$	$4.4965 \cdot 10^{-5}$
1	1	3	-2	$4.2336 \cdot 10^{-7}$	$-8.4868 \cdot 10^{-7}$
1	1	3	-1	$2.2679 \cdot 10^{-6}$	$-1.9140 \cdot 10^{-6}$
1	1	3	0	$7.3376 \cdot 10^{-7}$	$-1.0045 \cdot 10^{-6}$
1	1	3	1	$-3.7601 \cdot 10^{-7}$	$-2.8688 \cdot 10^{-7}$
1	1	3	2	$-9.2186 \cdot 10^{-8}$	$1.8182 \cdot 10^{-7}$
1	1	3	3	$2.2467 \cdot 10^{-8}$	$1.3965 \cdot 10^{-8}$
1	1	5	-5	$-2.2523 \cdot 10^{-6}$	$4.2296 \cdot 10^{-6}$
1	1	5	-4	$1.0777 \cdot 10^{-6}$	$-3.3309 \cdot 10^{-7}$
1	1	5	-3	$2.2294 \cdot 10^{-7}$	$-1.0088 \cdot 10^{-6}$
1	1	5	-2	$-1.7726 \cdot 10^{-7}$	$-5.9765 \cdot 10^{-9}$
1	1	5	-1	$-2.5302 \cdot 10^{-8}$	$8.2347 \cdot 10^{-8}$
1	1	5	0	$1.6505 \cdot 10^{-8}$	$8.9115 \cdot 10^{-9}$
1	1	5	1	$2.1029 \cdot 10^{-10}$	$-1.1379 \cdot 10^{-8}$
1	1	5	2	$-3.2058 \cdot 10^{-9}$	$-1.9943 \cdot 10^{-9}$
1	1	5	3	$2.4222 \cdot 10^{-9}$	$6.7578 \cdot 10^{-10}$
1	1	5	4	$1.4420 \cdot 10^{-10}$	$9.1725 \cdot 10^{-11}$
1	1	5	5	$-3.4628 \cdot 10^{-11}$	$-4.2841 \cdot 10^{-11}$

Input examples for the derivation of anisotropic exchange interaction are

given in listings S18 and S19. From the calculation with listing S18, the file `DyMn_vtzip_vdzip_cas_14in12_396r_aniso_cas.aniso` is obtained, which is used in listing S19 as `DATA`. The file `Dy_in_DyMn_vtzip_vdzip_cas_9in7_aniso_cas.old.aniso` is the data file for the single Dy site, where the Mn ion was computationally replaced by diamagnetic Zn, closed shell. The file `Mn_in_DyMn_vtzip_vdzip_cas_5in5_aniso_cas.old.aniso` is the data file for the single Mn site, where the Dy ion was computationally replaced by diamagnetic Lu, closed shell. The datafile for Mn can be in principle avoided in this particular case, but here it shown as an example for a general case.

Listing S18: Input file for the complete CASSCF/RASSI/SINGLE\_ANISO calculation for the entire binuclear DyMn compound.

```
1 &SEWARD
2 Title
3 DyMn full calculation
4 ANGM
5 2.21452658174807357384 15.28782765643874124177 15.91146562345711475748
6 AMFI
7
8 DOAN
9 RICD
10 SDIP
11
12 CDTH
13 1.0d-6
14
15 VERBOSE
16
17 Basis Set
18 DY.ANO-RCC-vtzip
19 DY1 1.17187700 8.08997000 8.41998500 Angstrom
20 End of Basis Set
21
22 Basis Set
23 MN.ANO-RCC-vdzip
24 MN2 1.50694300 4.68651000 6.93724100 Angstrom
25 End of Basis Set
26
27 Basis Set
28 O.ANO-RCC-vdzip
29 O3 1.00627200 12.53340600 9.80793300 Angstrom
30 End of Basis Set
31
32 Basis Set
33 O.ANO-RCC-vtzip
34 O4 1.22522100 9.93113900 9.56746000 Angstrom
35 End of Basis Set
36
37 Basis Set
38 O.ANO-RCC-vtzip
39 O5 1.75652500 5.89205800 8.74773800 Angstrom
40 End of Basis Set
41
42 Basis Set
43 O.ANO-RCC-vtzip
44 O6 1.60135100 9.51194200 6.50182000 Angstrom
```

```
45 End of Basis Set
46
47 Basis Set
48 O.ANO-RCC-vtzp
49 07 1.21159500 6.91777600 6.40048900 Angstrom
50 End of Basis Set
51
52 Basis Set
53 O.ANO-RCC-vdzp
54 08 1.00477500 2.53138800 6.26891000 Angstrom
55 End of Basis Set
56
57 Basis Set
58 O.ANO-RCC-vtzp
59 09 -0.98583600 6.97042900 8.84604400 Angstrom
60 End of Basis Set
61
62 Basis Set
63 O.ANO-RCC-vtzp
64 10 -0.99509200 8.80720300 7.66485800 Angstrom
65 End of Basis Set
66
67 Basis Set
68 O.ANO-RCC-vdzp
69 11 -0.63201900 4.60105000 7.59982400 Angstrom
70 End of Basis Set
71
72 Basis Set
73 O.ANO-RCC-vdzp
74 12 3.44899700 4.65167800 6.18421500 Angstrom
75 End of Basis Set
76
77 Basis Set
78 O.ANO-RCC-vtzp
79 13 3.54852900 8.26649900 8.32426500 Angstrom
80 End of Basis Set
81
82 Basis Set
83 N.ANO-RCC-vdzp
84 14 0.08557600 0.93155100 3.24107900 Angstrom
85 End of Basis Set
86
87 Basis Set
88 N.ANO-RCC-vdzp
89 15 0.14174200 -1.32847200 4.86540000 Angstrom
90 End of Basis Set
91
92 Basis Set
93 N.ANO-RCC-vtzp
94 16 1.58665300 7.46455500 10.75772200 Angstrom
95 End of Basis Set
96
97 Basis Set
98 N.ANO-RCC-vdzp
99 17 1.78989100 6.09153100 11.06927700 Angstrom
100 End of Basis Set
101
102 Basis Set
```

```

103 N.ANO-RCC-vdzp
104 N18 2.44356500 1.92993000 11.24925200 Angstrom
105 End of Basis Set
106
107 Basis Set
108 N.ANO-RCC-vdzp
109 N19 2.05104400 3.30700500 8.85360600 Angstrom
110 End of Basis Set
111
112 Basis Set
113 N.ANO-RCC-vdzp
114 N20 0.89677300 4.62332600 4.66576300 Angstrom
115 End of Basis Set
116
117 Basis Set
118 N.ANO-RCC-vdzp
119 N21 0.65267100 3.34143200 4.19843000 Angstrom
120 End of Basis Set
121
122 Basis Set
123 C.ANO-RCC-vdzp
124 C22 1.00291400 13.94895800 9.84725500 Angstrom
125 End of Basis Set
126
127 Basis Set
128 C.ANO-RCC-vdzp
129 C23 1.19756900 11.84486800 10.96189600 Angstrom
130 End of Basis Set
131
132 Basis Set
133 C.ANO-RCC-vdzp
134 C24 1.26631200 12.38354800 12.21265300 Angstrom
135 End of Basis Set
136
137 Basis Set
138 C.ANO-RCC-vdzp
139 C25 1.44510700 11.54515200 13.35754200 Angstrom
140 End of Basis Set
141
142 Basis Set
143 C.ANO-RCC-vdzp
144 C26 1.59162000 10.23693100 13.20781400 Angstrom
145 End of Basis Set
146
147 Basis Set
148 C.ANO-RCC-vdzp
149 C27 1.52604900 9.63952400 11.89504800 Angstrom
150 End of Basis Set
151
152 Basis Set
153 C.ANO-RCC-vdzp
154 C28 1.30542000 10.41716600 10.75620900 Angstrom
155 End of Basis Set
156
157 Basis Set
158 C.ANO-RCC-vdzp
159 C29 1.68129500 8.21789600 11.85572600 Angstrom
160 End of Basis Set

```

161  
162 Basis Set  
163 C.ANO-RCC-vdzp  
164 C30 1.85851900 5.40299300 9.95614800 Angstrom  
165 End of Basis Set  
166  
167 Basis Set  
168 C.ANO-RCC-vdzp  
169 C31 2.06444300 3.94288900 10.03479300 Angstrom  
170 End of Basis Set  
171  
172 Basis Set  
173 C.ANO-RCC-vdzp  
174 C32 2.27018200 3.25637700 11.20993000 Angstrom  
175 End of Basis Set  
176  
177 Basis Set  
178 C.ANO-RCC-vdzp  
179 C33 2.41049900 1.29809600 10.08319000 Angstrom  
180 End of Basis Set  
181  
182 Basis Set  
183 C.ANO-RCC-vdzp  
184 C34 2.23950300 1.98663300 8.87931700 Angstrom  
185 End of Basis Set  
186  
187 Basis Set  
188 C.ANO-RCC-vdzp  
189 C35 1.82739100 10.91736800 6.56382800 Angstrom  
190 End of Basis Set  
191  
192 Basis Set  
193 C.ANO-RCC-vdzp  
194 C36 1.43882700 8.94086100 5.20871500 Angstrom  
195 End of Basis Set  
196  
197 Basis Set  
198 C.ANO-RCC-vdzp  
199 C37 1.50073700 9.71242800 4.06231400 Angstrom  
200 End of Basis Set  
201  
202 Basis Set  
203 C.ANO-RCC-vdzp  
204 C38 1.34448300 9.07654300 2.85995400 Angstrom  
205 End of Basis Set  
206  
207 Basis Set  
208 C.ANO-RCC-vdzp  
209 C39 1.09956000 7.72579500 2.79643300 Angstrom  
210 End of Basis Set  
211  
212 Basis Set  
213 C.ANO-RCC-vdzp  
214 C40 1.06369800 6.95017800 3.97610700 Angstrom  
215 End of Basis Set  
216  
217 Basis Set  
218 C.ANO-RCC-vdzp

219 C41 1.23211900 7.59213700 5.20266600 Angstrom  
220 End of Basis Set  
221  
222 Basis Set  
223 C.ANO-RCC-vdzp  
224 C42 0.84336900 5.55892700 3.76890800 Angstrom  
225 End of Basis Set  
226  
227 Basis Set  
228 C.ANO-RCC-vdzp  
229 C43 0.71880100 2.34710200 5.07562400 Angstrom  
230 End of Basis Set  
231  
232 Basis Set  
233 C.ANO-RCC-vdzp  
234 C44 -0.18226100 -0.29364100 2.74198600 Angstrom  
235 End of Basis Set  
236  
237 Basis Set  
238 C.ANO-RCC-vdzp  
239 C45 0.42256600 -0.11745600 5.32063300 Angstrom  
240 End of Basis Set  
241  
242 Basis Set  
243 C.ANO-RCC-vtzip  
244 C46 -1.60812500 7.81489900 8.11706600 Angstrom  
245 End of Basis Set  
246  
247 Basis Set  
248 C.ANO-RCC-vdzp  
249 C47 -2.99249800 7.65694100 7.79341200 Angstrom  
250 End of Basis Set  
251  
252 Basis Set  
253 C.ANO-RCC-vdzp  
254 C48 4.68801500 4.84203800 6.91016900 Angstrom  
255 End of Basis Set  
256  
257 Basis Set  
258 C.ANO-RCC-vdzp  
259 C49 0.40420400 0.98825400 4.54325800 Angstrom  
260 End of Basis Set  
261  
262 Basis Set  
263 C.ANO-RCC-vdzp  
264 C50 -0.12635600 -1.40947700 3.55112200 Angstrom  
265 End of Basis Set  
266  
267 Basis Set  
268 H.ANO-RCC-vdzp  
269 H51 1.07519900 14.29322600 8.95493700 Angstrom  
270 End of Basis Set  
271  
272 Basis Set  
273 H.ANO-RCC-vdzp  
274 H52 1.74686200 14.25677400 10.37054600 Angstrom  
275 End of Basis Set  
276

```
277 Basis Set
278 H.ANO-RCC-vdzp
279 H53 0.18414300 14.25677400 10.24350500 Angstrom
280 End of Basis Set
281
282 Basis Set
283 H.ANO-RCC-vdzp
284 H54 1.19678700 13.30497300 12.32154600 Angstrom
285 End of Basis Set
286
287 Basis Set
288 H.ANO-RCC-vdzp
289 H55 1.45809500 11.91777200 14.20902500 Angstrom
290 End of Basis Set
291
292 Basis Set
293 H.ANO-RCC-vdzp
294 H56 1.73720100 9.69825200 13.95191600 Angstrom
295 End of Basis Set
296
297 Basis Set
298 H.ANO-RCC-vdzp
299 H57 1.86188100 7.78249800 12.65881200 Angstrom
300 End of Basis Set
301
302 Basis Set
303 H.ANO-RCC-vdzp
304 H58 2.29238800 3.73227800 12.00696600 Angstrom
305 End of Basis Set
306
307 Basis Set
308 H.ANO-RCC-vdzp
309 H59 2.50241500 0.37262000 10.06655300 Angstrom
310 End of Basis Set
311
312 Basis Set
313 H.ANO-RCC-vdzp
314 H60 2.25823600 1.51478200 8.07925600 Angstrom
315 End of Basis Set
316
317 Basis Set
318 H.ANO-RCC-vdzp
319 H61 2.55208100 11.15228100 5.97701600 Angstrom
320 End of Basis Set
321
322 Basis Set
323 H.ANO-RCC-vdzp
324 H62 2.05249800 11.16645700 7.46370800 Angstrom
325 End of Basis Set
326
327 Basis Set
328 H.ANO-RCC-vdzp
329 H63 1.03257800 11.38314300 6.29159600 Angstrom
330 End of Basis Set
331
332 Basis Set
333 H.ANO-RCC-vdzp
334 H64 1.64378000 10.62980200 4.10617400 Angstrom
```



```
335 End of Basis Set
336
337 Basis Set
338 H.ANO-RCC-vdzp
339 H65 1.40658000 9.56864500 2.07199200 Angstrom
340 End of Basis Set
341
342 Basis Set
343 H.ANO-RCC-vdzp
344 H66 0.95606900 7.32077300 1.97066100 Angstrom
345 End of Basis Set
346
347 Basis Set
348 H.ANO-RCC-vdzp
349 H67 0.64061600 5.29363800 2.90078900 Angstrom
350 End of Basis Set
351
352 Basis Set
353 H.ANO-RCC-vdzp
354 H68 -0.40505400 -0.38679600 1.84361900 Angstrom
355 End of Basis Set
356
357 Basis Set
358 H.ANO-RCC-vdzp
359 H69 0.64397800 -0.02227600 6.21900100 Angstrom
360 End of Basis Set
361
362 Basis Set
363 H.ANO-RCC-vdzp
364 H70 -3.07516000 7.26812000 6.92075500 Angstrom
365 End of Basis Set
366
367 Basis Set
368 H.ANO-RCC-vdzp
369 H71 -3.42323000 8.51558800 7.80248600 Angstrom
370 End of Basis Set
371
372 Basis Set
373 H.ANO-RCC-vdzp
374 H72 -3.40923400 7.08383500 8.44072000 Angstrom
375 End of Basis Set
376
377 Basis Set
378 H.ANO-RCC-vdzp
379 H73 5.40356200 4.40258900 6.44283600 Angstrom
380 End of Basis Set
381
382 Basis Set
383 H.ANO-RCC-vdzp
384 H74 4.60412400 4.46536800 7.79038700 Angstrom
385 End of Basis Set
386
387 Basis Set
388 H.ANO-RCC-vdzp
389 H75 4.87867600 5.77966400 6.97973900 Angstrom
390 End of Basis Set
391
392 Basis Set
```

```

393 H.ANO-RCC-vdzp
394 H76 -0.28178900 -2.24584700 3.17604600 Angstrom
395 End of Basis Set
396
397 Basis Set
398 H.ANO-RCC-vdzp
399 H77 -1.13276100 4.43499100 6.93134200 Angstrom
400 End of Basis Set
401
402 Basis Set
403 H.ANO-RCC-vdzp
404 H78 -0.85063100 5.35236600 7.93709000 Angstrom
405 End of Basis Set
406
407 Basis Set
408 H.ANO-RCC-vdzp
409 H79 3.63573900 4.41474000 5.38566700 Angstrom
410 End of Basis Set
411
412 Basis Set
413 H.ANO-RCC-vtzp
414 H80 4.09218800 7.94045600 8.89444100 Angstrom
415 End of Basis Set
416
417 Basis Set
418 H.ANO-RCC-vtzp
419 H81 3.98627600 8.66949600 7.71325500 Angstrom
420 End of Basis Set
421
422 End of Input
423
424 &RASSCF
425 SYMM
426 1
427 SPIN
428 11
429 NACTEL
430 14 0 0
431 FROZ
432 0
433 INAC
434 213
435 RAS1
436 0
437 RAS2
438 12
439 RAS3
440 0
441 CIROOT
442 11 11 1
443 CIMX
444 25
445 ORBL
446 ALL
447 ORBA
448 FULL
449 End of Input
450

```

```

451 >>> COPY $WorkDir/$Project.RasOrb
    ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms11_11r.RasOrb
452 >>> COPY $WorkDir/$Project.RunFile
    ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms11_11r.RunFile
453 >>> COPY $WorkDir/$Project.rasscf.h5
    ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms11_11r.rasscf.h5
454 >>> COPY $WorkDir/$Project.JobIph
    ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms11_11r.JobIph
455
456 >>COPY $Project.RasOrb INPORB
457 &RASSCF
458     LUMO
459     SYMM
460         1
461     SPIN
462         9
463     NACTEL
464         14 0 0
465     FROZ
466         0
467     INAC
468         213
469     RAS1
470         0
471     RAS2
472         12
473     RAS3
474         0
475     CIROOT
476         11 11 1
477     CIMX
478         25
479     ORBL
480     ALL
481     ORBA
482     FULL
483 End of Input
484
485 >>> COPY $WorkDir/$Project.RasOrb
    ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms9_11r.RasOrb
486 >>> COPY $WorkDir/$Project.RunFile
    ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms9_11r.RunFile
487 >>> COPY $WorkDir/$Project.rasscf.h5
    ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms9_11r.rasscf.h5
488 >>> COPY $WorkDir/$Project.JobIph
    ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms9_11r.JobIph
489
490 &RASSCF
491     LUMO
492     SYMM
493         1
494     SPIN
495         7
496     NACTEL
497         14 0 0
498     FROZ
499         0
500     INAC

```

```

501      213
502  RAS1
503      0
504  RAS2
505      12
506  RAS3
507      0
508  CIROOT
509      11 11 1
510  CIMX
511      25
512  ORBL
513  ALL
514  ORBA
515  FULL
516 End of Input
517
518 >>> COPY $WorkDir/$Project.RasOrb
519 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms7_11r.RasOrb
520 >>> COPY $WorkDir/$Project.RunFile
521 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms7_11r.RunFile
522 >>> COPY $WorkDir/$Project.rasscf.h5
523 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms7_11r.rasscf.h5
524 >>> COPY $WorkDir/$Project.JobIph
525 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms7_11r.JobIph
526
527 &RASSCF
528 LUMO
529 SYMM
530      1
531 SPIN
532      5
533 NACTEL
534      14 0 0
535 FROZ
536      0
537 INAC
538      213
539 RAS1
540      0
541 RAS2
542      12
543 RAS3
544      0
545 CIROOT
546      11 11 1
547 CIMX
548      25
549 ORBL
550 ALL
551 ORBA
552 FULL
553 End of Input
554
555 >>> COPY $WorkDir/$Project.RasOrb
556 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms5_11r.RasOrb
557 >>> COPY $WorkDir/$Project.RunFile

```

```

554 ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms5_11r.RunFile
>>> COPY $WorkDir/$Project.rasscf.h5
555 ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms5_11r.rasscf.h5
>>> COPY $WorkDir/$Project.JobIph
556 ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms5_11r.JobIph
557
558 &RASSCF
559     LUMO
560     SYMM
561         1
562     SPIN
563         3
564     NACTEL
565         14 0 0
566     FROZ
567         0
568     INAC
569         213
570     RAS1
571         0
572     RAS2
573         12
574     RAS3
575         0
576     CIROOT
577         11 11 1
578     CIMX
579         25
580     ORBL
581     ALL
582     ORBA
583     FULL
584 End of Input
585
586 >>> COPY $WorkDir/$Project.RasOrb
587 ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms3_11r.RasOrb
>>> COPY $WorkDir/$Project.RunFile
588 ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms3_11r.RunFile
>>> COPY $WorkDir/$Project.rasscf.h5
589 ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms3_11r.rasscf.h5
>>> COPY $WorkDir/$Project.JobIph
590 ↪ $FileDir/DyMn_vtzip_vdzip_cas_14in12_rasscf_ms3_11r.JobIph
591
592 &RASSCF
593     LUMO
594     SYMM
595         1
596     SPIN
597         1
598     NACTEL
599         14 0 0
600     FROZ
601         0
602     INAC
603         213
604     RAS1

```

```

605      0
606      RAS2
607      12
608      RAS3
609      0
610      CIROOT
611      11 11 1
612      CIMX
613      25
614      ORBL
615      ALL
616      ORBA
617      FULL
618      End of Input
619
620      >>> COPY $WorkDir/$Project.RasOrb
621      ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms1_11r.RasOrb
622      >>> COPY $WorkDir/$Project.RunFile
623      ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms1_11r.RunFile
624      >>> COPY $WorkDir/$Project.rasscf.h5
625      ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms1_11r.rasscf.h5
626      >>> COPY $WorkDir/$Project.JobIph
627      ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms1_11r.JobIph
628
629      >>> COPY $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms11_11r.JobIph $WorkDir/JOB001
630      >>> COPY $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms9_11r.JobIph $WorkDir/JOB002
631      >>> COPY $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms7_11r.JobIph $WorkDir/JOB003
632      >>> COPY $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms5_11r.JobIph $WorkDir/JOB004
633      >>> COPY $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms3_11r.JobIph $WorkDir/JOB005
634      >>> COPY $FileDir/DyMn_vtzp_vdzp_cas_14in12_rasscf_ms1_11r.JobIph $WorkDir/JOB006
635
636      &RASSI
637      OMEG
638      J-VA
639      MEES
640      Prop
641      6
642      AngMom 1
643      AngMom 2
644      AngMom 3
645      MLTPL1 1
646      MLTPL1 2
647      MLTPL1 3
648      SpinOrbit
649      NR OF JOBIPHS
650      6 11 11 11 11 11 11
651      1 2 3 4 5 6 7 8 9 10 11
652      1 2 3 4 5 6 7 8 9 10 11
653      1 2 3 4 5 6 7 8 9 10 11
654      1 2 3 4 5 6 7 8 9 10 11
655      1 2 3 4 5 6 7 8 9 10 11
656      EJOB
657      End of Input
658      >>> COPY $WorkDir/$Project.rassi.h5

```

```

659 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_rassi_cas.h5
>>> COPY $WorkDir/$Project.RunFile
660 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_rassi_cas.RunFile
661 &SINGLE_ANISO
662 MLTP
663 7
664 2 2 2 2 2 2 2
665 PLOT
666 TINT
667 0 300 301
668 XFIE
669 0.50
670 HINT
671 0 7.0 71
672 TMAG
673 5 2.0 3.0 5.0 8.0 10.0
674 MAVE
675 1 10
676 UBAR
677 End Of Input
678 >>> COPY $WorkDir/$Project.aniso
679 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas.aniso
>>> COPY $WorkDir/$Project.MH.dat
680 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_MH.dat
>>> COPY $WorkDir/$Project.MH.plt
681 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_MH.plt
>>> COPY $WorkDir/$Project.MH.png
682 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_MH.png
>>> COPY $WorkDir/$Project.XT_no_field.dat
683 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_XT_no_field.dat
>>> COPY $WorkDir/$Project.XT_no_field.plt
684 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_XT_no_field.plt
>>> COPY $WorkDir/$Project.XT_no_field.png
685 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_XT_no_field.png
>>> COPY $WorkDir/$Project.XT_with_field_dM_over_dH.dat
686 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_XT_with_field_dM_over_dH.dat
>>> COPY $WorkDir/$Project.XT_with_field_dM_over_dH.plt
687 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_XT_with_field_dM_over_dH.plt
>>> COPY $WorkDir/$Project.XT_with_field_dM_over_dH.png
688 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_XT_with_field_dM_over_dH.png
>>> COPY $WorkDir/$Project.XT_with_field_M_over_H.dat
689 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_XT_with_field_M_over_H.dat
>>> COPY $WorkDir/$Project.XT_with_field_M_over_H.plt
690 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_XT_with_field_M_over_H.plt
>>> COPY $WorkDir/$Project.XT_with_field_M_over_H.png
691 ↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas_XT_with_field_M_over_H.png
>>> COPY $WorkDir/$Project.old.aniso
↪ $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas.old.aniso

```

Listing S19: Input file showing the function *PREX*.

```

1 >>>COPY $FileDir/DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas.aniso $WorkDir/
2 >>>COPY $FileDir/Dy_in_DyMn_vtzp_vdzp_cas_9in7_aniso_cas.old.aniso $WorkDir/
3 >>>COPY $FileDir/Mn_in_DyMn_vtzp_vdzp_cas_5in5_aniso_cas.old.aniso $WorkDir/
4
5 &SINGLE_ANISO
6 DATA

```

```

7 DyMn_vtzp_vdzp_cas_14in12_396r_aniso_cas.aniso
8 PREX
9 2 Dy_in_DyMn_vtzp_vdzp_cas_9in7_aniso_cas.old.aniso
10 6 Mn_in_DyMn_vtzp_vdzp_cas_5in5_aniso_cas.old.aniso
11 TINT
12 0 300 31
13 MLTP
14 6
15 2 2 2 2 2 2
16 End Of Input

```

## S6 Relativistic Hyperfine Coupling

In listing S20 – a part of the verification test suite, `test/extra/396.input` – an input sample to perform a hyperfine coupling calculation using the GEN1INT library for the relevant one-electron integrals is given. To activate the hyperfine coupling features, OpenMolcas must be compiled with the GEN1INT option (e.g., `cmake -D GEN1INT=ON`). In the OpenMolcas SEWARD module, the `FLDG=0` keyword indicates that hyperfine coupling constants will be calculated for each nucleus, `AMFI` requests the computation of the atomic mean-field integrals for the inclusion of spin-orbit coupling effects in the RASSI module, `MXTC` indicates the calculation of X2C transformed hyperfine magnetic integrals to be used in the hyperfine calculations, note that `MXTC` must be used together with the `RX2C` keyword, and finally `clight` specifies the speed of light. For SCF or DFT calculations, the `HFC` keyword activates the hyperfine coupling tensor matrix calculation on each atom, and the `HFC` keyword must be used in conjunction with UHF.

With the RASSI module, the following keywords are required: `SPIN` requests the spin-orbit interaction matrix elements and `EPRA` computes the hyperfine tensor matrix and principal magnetic axes values for the ground spin-orbit state. For the hyperfine matrix elements, calculating both spin-dependent (ASD) and the paramagnetic spin-orbital (PSOP) parts is recommended. ASD and PSOS are calculated using the `PROP` keyword. Note that there are 6 ASD and 3 PSOP components for each atom. The following keywords are optional: `AFCC` gives separate Fermi contact contributions, `ASDC` the spin-dipolar contributions, `FCSD` the total spin-dependent contributions (AFCC+ASDC), `APSO` computes the paramagnetic spin-orbital contributions independently. Note that keywords `AFCC`, `ASDC`, and `FCSD` require the ASD part of the hyperfine coupling matrix, while `APSO` requires the PSOP part. Finally, `ATSA` activates the pseudospin approach to calculate the hyperfine constants for non-doublet states. If the `ATSA` keyword is not present, but a state has more than two components, the code will exit with an error. If the ground state is a Kramers doublet, the pseudospin approach will result in the same hyperfine coupling as the standard approach.

Listing S20: OpenMolcas input for an HFC calculation.

```

1 &SEWARD
2 FINite
3 cholesky
4 BSSHOW
5 Basis set
6 Si.cc-pVTZ-DK..15s9p2d1f.5s4p2d1f.
7 Si      1.641895    -0.007260    0.000000    angstrom
8 End of basis

```



```

9 Basis set
10 S.cc-pVTZ-DK..15s9p2d1f.5s4p2d1f.
11 S      -0.486365   -0.002825   0.000000   angstrom
12 end of basis
13 Basis set
14 H.cc-pVTZ-DK..5s2p1d.3s2p1d.
15 H      -0.758580   1.333849   0.000000   angstrom
16 end of basis
17 RX2C
18 clight
19 137.036
20 Angmom
21   0.0 0.0 0.0
22 FLDG
23   0
24 AMFI
25 MXTC
26 End of input
27
28 &SCF
29 UHF
30 HFC
31 THREsholds
32 1.0d-10 1.0d-8 1.5d-4 0.2d-4
33 charge=0
34 End of input
35
36 &RASSCF
37 Title= The CASSCF of NpF6 with an active space of (1,8)
38 Spin
39 2
40 Symmetry
41 1
42 nActEl
43 1 0 0
44 Inactive
45 15
46 Ras2
47 8
48 CIRoots
49 8 8 1
50 Iter
51 100 100
52 OrbAppear
53 Compact
54 LumOrb
55 levshift
56 1.0
57 End of input
58
59 &RASSI
60 NrofJobIphs
61 1 8
62 1 2 3 4 5 6 7 8
63 SPIN
64 EPRA
65 AFCC
66 ASDC

```

```

67 FCSD
68 APSO
69 ATSA
70 PROPERTIES
71 27
72 'ASD 1' 1
73 'ASD 1' 2
74 'ASD 1' 3
75 'ASD 1' 4
76 'ASD 1' 5
77 'ASD 1' 6
78 'ASD 2' 1
79 'ASD 2' 2
80 'ASD 2' 3
81 'ASD 2' 4
82 'ASD 2' 5
83 'ASD 2' 6
84 'ASD 3' 1
85 'ASD 3' 2
86 'ASD 3' 3
87 'ASD 3' 4
88 'ASD 3' 5
89 'ASD 3' 6
90 'PSOP 1' 1
91 'PSOP 1' 2
92 'PSOP 1' 3
93 'PSOP 2' 1
94 'PSOP 2' 2
95 'PSOP 2' 3
96 'PSOP 3' 1
97 'PSOP 3' 2
98 'PSOP 3' 3
99 end of input

```

## S7 Single and Two-Photon Spectra with the Exact Semi-Classical Operator

Input files for performing photon-in photon-out calculations using the exact semi-classical light-matter operator. Notice that the inputs files are for a RASPT2 calculations, but this methodology works with RASSCF too.

Listing S21: tpu.xyz: Fe<sup>II</sup>(P)(ImH)<sub>2</sub> geometry.

```

1 55
2 FeII(P)(ImH)2 B3LYP/6-32G(d) geometry
3 FE 0.00000000 0.00000000 0.00000000
4 N -1.41871810 1.42829924 -0.05008112
5 C -2.77860804 1.23671423 -0.04241476
6 C -3.47045523 2.50649453 -0.08145791
7 H -4.54733899 2.62348940 -0.08542147
8 C -2.51334757 3.47244019 -0.11097290
9 H -2.63967034 4.54784164 -0.14500326
10 C -1.23755992 2.79042795 -0.09118545
11 C 0.00000000 3.42986011 -0.10875308
12 H 0.00000000 4.51616347 -0.14255035

```

13	C	3.41304412	0.00000000	0.00000000
14	H	4.49975735	0.00000000	0.00000000
15	N	0.00000000	0.07346165	2.01412485
16	C	0.00000000	1.16913437	2.75418232
17	H	0.00000000	2.18010867	2.37849152
18	N	0.00000000	0.85276984	4.07481867
19	H	0.00000000	1.51393411	4.83664673
20	C	0.00000000	-0.52475535	4.18191792
21	H	0.00000000	-1.02921946	5.13541076
22	C	0.00000000	-0.99044340	2.89607911
23	H	0.00000000	-2.00856863	2.54191494
24	N	-1.41871810	-1.42829924	0.05008112
25	N	1.41871810	-1.42829924	0.05008112
26	N	1.41871810	1.42829924	-0.05008112
27	C	-2.77860804	-1.23671423	0.04241476
28	C	2.77860804	-1.23671423	0.04241476
29	C	2.77860804	1.23671423	-0.04241476
30	C	-3.47045523	-2.50649453	0.08145791
31	C	3.47045523	-2.50649453	0.08145791
32	C	3.47045523	2.50649453	-0.08145791
33	H	-4.54733899	-2.62348940	0.08542147
34	H	4.54733899	-2.62348940	0.08542147
35	H	4.54733899	2.62348940	-0.08542147
36	C	-2.51334757	-3.47244019	0.11097290
37	C	2.51334757	-3.47244019	0.11097290
38	C	2.51334757	3.47244019	-0.11097290
39	H	-2.63967034	-4.54784164	0.14500326
40	H	2.63967034	-4.54784164	0.14500326
41	H	2.63967034	4.54784164	-0.14500326
42	C	-1.23755992	-2.79042795	0.09118545
43	C	1.23755992	-2.79042795	0.09118545
44	C	1.23755992	2.79042795	-0.09118545
45	C	0.00000000	-3.42986011	0.10875308
46	H	0.00000000	-4.51616347	0.14255035
47	C	-3.41304412	0.00000000	0.00000000
48	H	-4.49975735	0.00000000	0.00000000
49	N	0.00000000	-0.07346165	-2.01412485
50	C	0.00000000	-1.16913437	-2.75418232
51	H	0.00000000	-2.18010867	-2.37849152
52	N	0.00000000	-0.85276984	-4.07481867
53	H	0.00000000	-1.51393411	-4.83664673
54	C	0.00000000	0.52475535	-4.18191792
55	H	0.00000000	1.02921946	-5.13541076
56	C	0.00000000	0.99044340	-2.89607911
57	H	0.00000000	2.00856863	-2.54191494

Listing S22: fe2-ce.input: Input for computing the ground state and 1s and 2p core-excitations. It calls for a pre-computed reference orbital file (ref.RasOrb).

```

1  *-----
2  * fe2-ce.input - Compute the ground state and 1s and 2p core-excitations
3  * .                and save the JobMix files.
4  *-----
5  &GATEWAY
6  TITLE = "RIXS map of FeII(P)(ImH)2 using the exact operator"
7  COORD = tpu.xyz
8  Basis = ANO-RCC-VDZP
9  RICD

```

```

10 &SEWARD
11 *-----
12 ** Ground state
13 *-----
14 >>> COPY $CurrDir/ref.RasOrb INPORB
15 &RASSCF
16 LUMORB
17 SYMMETRY
18 1
19 Spin
20 1
21 Charge
22 0
23 RASSCF
24 1 2
25 SupSym
26 1 ; 1 45
27 0 ;
28 1 ; 1 25
29 1 ; 2 39 40
30 CIR0ot
31 1 1 1
32 >>> COPY $Project.RasOrb $CurrDir/fe2-gs.RasOrb
33 *-----
34 &CASPT2
35 MULTistate
36 All
37 Imaginary
38 0.3
39 MAXIter
40 200
41 >>> COPY $Project.JobMix $CurrDir/$Project.gs1.JobMix
42 *-----
43 ** 1s core-excitation
44 *-----
45 >>> foreach SYMMETRY in ( 1, 2 )
46 >>> COPY $CurrDir/fe2-gs.RasOrb INPORB
47 &RASSCF
48 LUMORB
49 SYMMETRY
50 $SYMMETRY
51 Spin
52 1
53 Charge
54 0
55 RASSCF
56 1 2
57 SUPSYM
58 1 ; 1 45
59 0 ;
60 1 ; 1 25
61 1 ; 2 39 40
62 HEXS
63 1
64 3
65 CIR0ot
66 60 60 1
67 *-----

```

```

68 &CASPT2
69 MULTistate
70 All
71 Imaginary
72 0.3
73 MAXIter
74 200
75 >>> COPY $Project.JobMix $CurrDir/$Project.1s.s$SYMMETRY.$SPIN.JobMix
76 >>> end do
77 *-----
78 ** 2p core-excitation
79 *-----
80 >>> foreach SYMMETRY in ( 3, 4 )
81 >>> foreach SPIN in ( 1, 3 )
82 >>> COPY $CurrDir/fe2-gs.RasOrb INPORB
83 &RASSCF
84 LUMORB
85 SYMMETRY
86 3
87 Spin
88 1
89 Charge
90 0
91 RASSCF
92 1 2
93 SUPSYM
94 1 ; 1 45
95 0 ;
96 1 ; 1 25
97 1 ; 2 39 40
98 CIROot
99 120 120 1
100 *-----
101 &CASPT2
102 MULTistate
103 All
104 Imaginary
105 0.3
106 MAXIter
107 200
108 >>> COPY $Project.JobMix $CurrDir/$Project.2p.s$SYMMETRY.$SPIN.JobMix
109 >>> end do
110 >>> end do
111 *-----

```

Listing S23: `fe2-rixs-exact.input`: Input for computing the RIXS maps using the full exact operator. It calls for previously computed ground and core-excited states `JobMix` files as described above.

```

1 *-----
2 * fe2-rixs-exact.input - Compute the RIXS map with the full exact operator.
3 *-----
4 &GATEWAY
5 TITLE = "RIXS map of FeII(P)(ImH)2 using the exact operator"
6 COORD = tpu.xyz
7 Basis = ANO-RCC-VDZP
8 RICD
9 angmom

```

```

10 0.00 0.00 0.00
11 omqi
12 0.00 0.00 0.00
13 *-----
14 &SEWARD
15 multipole
16 3
17 center
18 3
19 1 0.0 0.0 0.0
20 2 0.0 0.0 0.0
21 3 0.0 0.0 0.0
22 *-----
23 >>> COPY $CurrDir/fe2-ce.gs.JobMix JOB001
24 >>> COPY $CurrDir/fe2-ce.2p.s3.1.JobMix JOB002
25 >>> COPY $CurrDir/fe2-ce.2p.s4.1.JobMix JOB003
26 >>> COPY $CurrDir/fe2-ce.2p.s3.3.JobMix JOB004
27 >>> COPY $CurrDir/fe2-ce.2p.s4.3.JobMix JOB005
28 >>> COPY $CurrDir/fe2-ce.1s.s1.JobMix JOB006
29 >>> COPY $CurrDir/fe2-ce.1s.s2.JobMix JOB007
30 *-----
31 &RASSI &END
32 Ejob
33 NROF jobiphs
34 7 All
35 SUBS
36 481
37 TINT
38 QIALL
39 qipr
40 1e-8
41 SPINorbit
42 *-----

```

Listing S24: fe2-rixs-exact-group.input: Input for computing the RIXS maps using the exact operator with the grouping scheme.

```

1 *-----
2 * fe2-rixs-exact-group.input -
3 * Compute the RIXS map with the exact operator and grouping scheme.
4 *-----
5 &GATEWAY
6 TITLE = "RIXS map of FeII(P)(ImH)2 using the exact operator"
7 COORD = tpu.xyz
8 Basis = ANO-RCC-VDZP
9 RICD
10 angmom
11 0.00 0.00 0.00
12 omqi
13 0.00 0.00 0.00
14 *-----
15 &SEWARD
16 multipole
17 3
18 center
19 3
20 1 0.0 0.0 0.0
21 2 0.0 0.0 0.0

```

```

22 3 0.0 0.0 0.0
23 *-----
24 >>> COPY $CurrDir/fe2-ce.gs.JobMix      JOB001
25 >>> COPY $CurrDir/fe2-ce.2p.s3.1.JobMix  JOB002
26 >>> COPY $CurrDir/fe2-ce.2p.s4.1.JobMix  JOB003
27 >>> COPY $CurrDir/fe2-ce.2p.s3.3.JobMix  JOB004
28 >>> COPY $CurrDir/fe2-ce.2p.s4.3.JobMix  JOB005
29 >>> COPY $CurrDir/fe2-ce.1s.s1.JobMix    JOB006
30 >>> COPY $CurrDir/fe2-ce.1s.s2.JobMix    JOB007
31 *-----
32 &RASSI &END
33 Ejob
34 NROF jobiphs
35 7 All
36 SUBS
37 481
38 TINT
39 TIGr
40 0.001
41 QIALL
42 qipr
43 1e-8
44 SPINorbit
45 *-----

```

## S8 Vibrational–Rotational Levels for Diatomics

An input example for a LEVEL calculation is given in listing S25, and an excerpt of the output in listing S26.

Listing S25: Input example for a LEVEL calculation.

```

1 &LEVEL
2
3 IAN1 = 3 # Atomic number for atom 1
4 IMN1 = 6 # Mass number for atom 1
5 IAN2 = 3 # Atomic number for atom 2
6 IMN2 = 6 # Mass number for atom 2
7 CHARGE = 0 # Charge of the molecule
8 VLIM = 0.0d0 # Energy at dissociation (cm-1)
9 RH = 0.0005 # ODE step size
10 RMIN = 0.125 # ODE lower limit
11 PPAR = 5 # p
12 QPAR = 3 # q
13 MMLR = 6 8 10 # Long-range inverse powers
14 DSCM = 333.768 # Depth at equilibrium (cm-1)
15 REQ = 4.17001 # Length at equilibrium (Angstrom)
16 RREF = 8.0d0 # Reference distance (Angstrom)
17 CMM = 6.719e6 1.12635e8 2.78694e9 # Long-range constants (C6,C8,C10)
18 PARM = -5.16e-1 -9.59e-2 1.17e-1 -2.28e-2 # Beta 0, Beta 1, Beta 2, Beta 3
19 LCDC = 0

```

Listing S26: Output excerpt for a LEVEL calculation.

```

1 SUMMARY (ALL ENERGIES IN CM-1):
2

```

Find 10 Potential-1 vibrational levels with J= 0								
	v	E(v)	v	E(v)	v	E(v)	v	E(v)
	0	-299.3817	3	-133.8550	6	-34.6437	9	-0.8188
	1	-236.6922	4	-93.4264	7	-16.4357		
	2	-181.5854	5	-60.3267	8	-5.4964		
An	n= 6 N-D theory extrapolation from last 2 levels implies						vD =	10.128

## S9 Evaluating Frenkel's Excitonic Coupling Terms

Computational details and results for the applications discussed in section 3.6.

### S9.1 Azulene

#### S9.1.1 Monomer geometries and energetics

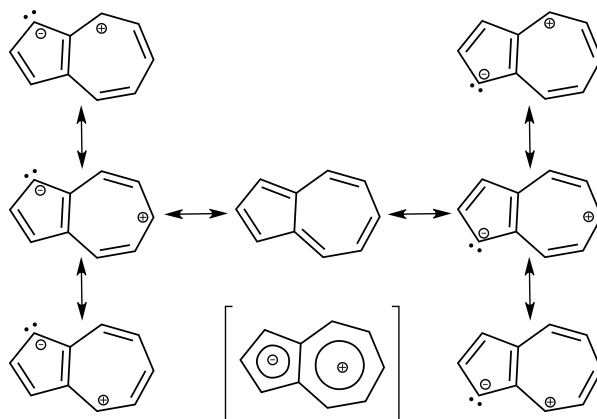


Figure S4: **Azulene molecule.** A few of the azulene resonance structures are shown here, together with the corresponding resonance hybrid (bottom middle). Only those formulas with charge separation imply an aromatic compound, hence they are the most important ones. Aromaticity induces a dipole from the 7-membered to the 5-membered ring.

Azulene is an aromatic molecule, constituted by five-membered and seven-membered fused rings, with structure as shown in fig. S4. Azulene electronic excitation spectrum is characterized by both  $\pi \rightarrow \pi^*$  and  $\pi \rightarrow$  Rydberg transitions. Thus, an extra basis set, specifically tailored for the azulene molecule, was created using the module **GENANO**, starting from a general Rydberg basis set.<sup>S7</sup> The obtained basis set (which we called **AZURYDBERG**, listing S27) provides one basis function for each s,  $p_x$ ,  $p_y$ ,  $p_z$  orbitals, and is applied to a ghost atom (X) located on the molecule centre of mass.

The geometry of azulene was optimized in vacuo at the SA-CASSCF level of theory,<sup>S8</sup> using the triple zeta basis ANO-RCC-VTZP,<sup>S9</sup> plus **AZURYDBERG**. State averaging was done over 8 states to preserve the Rydberg virtual orbitals inside the active space, although the optimization was, obviously, performed on the ground state. The active space comprised 10 electrons and 14 RAS2 orbitals (table S4), of which there were 5  $\pi$ , 5  $\pi^*$ , and 4 Rydberg virtual (s-type,  $p_{x,y,z}$ -type). We used a level shift of  $0.1 E_h$  to help wave function optimization convergence.



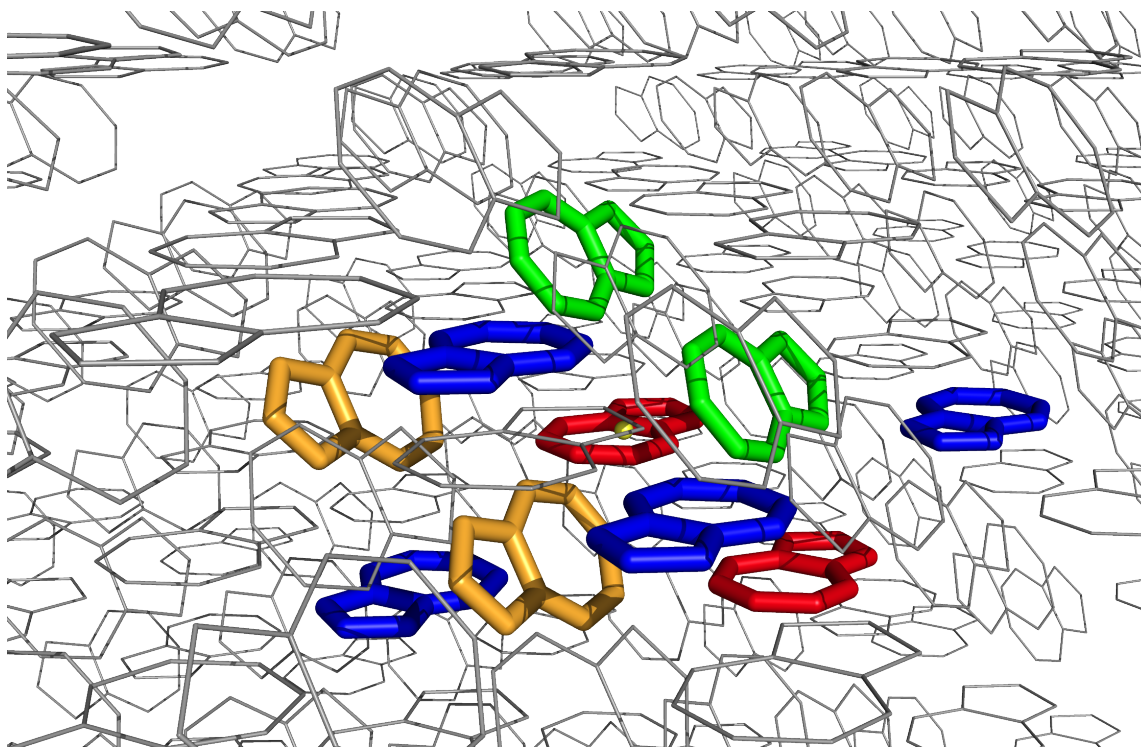


Figure S5: **Azulene crystal representation.** Among the 500 molecules azulene  $5 \times 5 \times 5$  supercell (grey lines, only a fraction is shown here), we highlighted four different types, which we refer to as: flat (red sticks), inverted flat (blue sticks), tilted (green sticks), and inverted titled (orange sticks). Hydrogen atoms are hidden for the sake of readability. We defined all possible combinations of unique dimers as those found within  $25 \text{ \AA}$  from the flat azulene monomer in the middle of the figure, highlighted by a yellow dot. These sets of dimers were later used to compute excitonic couplings, to be extended to the rest of the crystal.

The optimized geometry (listing S28) corresponded to the reference “flat” geometry highlighted in fig. S5. Azulene monomers inside the crystal structure<sup>S10</sup> can have four different orientations, which we refer to as “flat”, “inverted flat”, “tilted”, and “inverted tilted”, as shown in fig. S5. Thus, the flat optimized geometry was rigidly projected onto the other three different orientations (listings S29 to S31).

Single point energy evaluations were performed for the four monomer orientations, at the same level of theory as the previous geometry optimization (8roots-SA-CASSCF, active space with 10 electrons in 14 orbitals). Energies were subsequently re-evaluated at the MS-CASPT2 level of theory,<sup>S11</sup> with default IPEA shift value  $0.25 E_h$ ,<sup>S12</sup> and an imaginary level shift of  $0.1 E_h$ .<sup>S13</sup> Computed energetics for the azulene monomer are reported in table S5. All calculation used Cholesky decomposition for speed up,<sup>S14-S16</sup> and no symmetry constraint.

Listing S27: Rydberg extra basis functions specifically tailored for azulene.

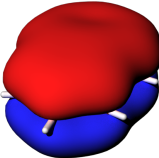
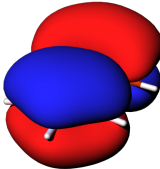

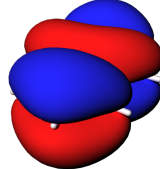
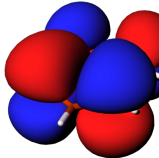
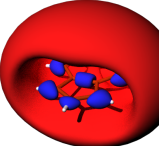
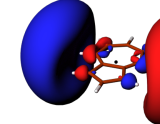
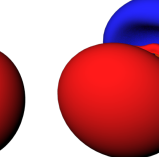
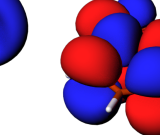
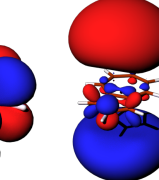
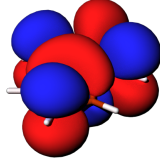

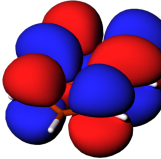

```

1 /X.AZURYDBERG..8s8p8d.1s1p1d.
2 Rydberg orbitals for Azulene
3 GHOST (8s,8p,8d) -> [1s,1p,1d]
4     0.0  0
5 * s-type functions
6     8    1
7     0.02462393256690
8     0.01125334281667
9     0.00585838058488
10    0.00334597391176
11    0.00204842249316
12    0.00132364236431
13    0.00089309575944
14    0.00062431294198
15     0.3776806518
16    -1.1846097118
17     0.1486302928
18    -0.6983559229
19     0.8163346267
20    -0.6988936683
21     0.3755558467
22    -0.0930378890
23 * p-type functions
24     8    1
25     0.04233528041629
26     0.01925420565070
27     0.00998821043479
28     0.00568936060666
29     0.00347567961959
30     0.00224205895181
31     0.00151063985318
32     0.00105475270898
33     0.1524651841
34    -0.0598750355
35    -0.8289270600
36    -0.0490334535
37    -0.3174558759
38     0.2469509609
39    -0.1398163117
40     0.0370466341
41 * d-type functions
42     8    1

```

43	0.06054020300665
44	0.02744569273592
45	0.01420439913628
46	0.00807659325172
47	0.00492718647974
48	0.00317481115512
49	0.00213712310185
50	0.00149101554039
51	0.0557392505
52	-0.0066265130
53	-0.3854752598
54	-0.3837364797
55	-0.3078424907
56	0.0203538284
57	-0.0471536110
58	0.0125574710

Table S4: Active space orbitals for azulene. Electronic occupation for each state is given under each orbital ( $S_0$ – $S_7$  from top to bottom). Density contour value = 0.0121.

					
$S_0$	1.958	1.939	1.934	1.827	1.820
$S_1$	1.944	1.913	1.874	1.786	1.108
$S_2$	1.954	1.926	1.901	1.866	1.005
$S_3$	1.938	1.917	1.837	1.533	1.368
$S_4$	1.954	1.926	1.900	1.867	1.007
$S_5$	1.931	1.896	1.691	1.390	1.556
$S_6$	1.954	1.926	1.902	1.863	1.004
$S_7$	1.947	1.923	1.851	1.790	1.102
					
$S_0$	0.000	0.000	0.000	0.047	0.000
$S_1$	0.000	0.000	0.000	0.059	0.001
$S_2$	0.989	0.005	0.006	0.032	0.000
$S_3$	0.000	0.000	0.000	0.055	0.002
$S_4$	0.006	0.993	0.001	0.031	0.000
$S_5$	0.000	0.000	0.000	0.061	0.173
$S_6$	0.009	0.001	0.990	0.032	0.000
$S_7$	0.000	0.000	0.000	0.038	0.858
					
$S_0$	0.183	0.063	0.184	0.045	
$S_1$	0.908	0.094	0.247	0.065	
$S_2$	0.125	0.060	0.095	0.037	
$S_3$	0.585	0.098	0.595	0.072	
$S_4$	0.123	0.059	0.094	0.038	
$S_5$	0.631	0.084	0.521	0.065	
$S_6$	0.127	0.058	0.096	0.038	
$S_7$	0.227	0.066	0.152	0.046	

Listing S28: Atom coordinates for the optimized flat geometry (in Å). X is a label for the center of mass of the molecule.

1	C1	15.0157279	1.1938718	9.1716026
2	C2	14.7261655	0.2530861	10.1694956
3	C3	14.0281222	-0.8266259	9.6117327
4	C4	13.2166484	-1.4309588	7.3449705
5	C5	13.0143820	-1.2561629	5.9744097
6	C6	13.4083353	-0.1933988	5.1607788
7	C7	14.1038444	0.9708566	5.4892623
8	C8	14.5905183	1.3781646	6.7329156
9	C9	14.5039863	0.7267595	7.9545605
10	C10	13.8599947	-0.5905520	8.2414124
11	H1	15.5391949	2.1168250	9.3072465
12	H2	14.9970753	0.3454082	11.2013624
13	H3	13.6779839	-1.6905151	10.1364765
14	H4	12.8205927	-2.3388077	7.7634050
15	H5	12.4835158	-2.0469373	5.4789478
16	H6	13.1353647	-0.2853835	4.1256077
17	H7	14.2896664	1.6449384	4.6745772
18	H8	15.1012124	2.3242710	6.7475135
19	X	14.0398182	0.0227078	7.5707620

Listing S29: Atom coordinates for the optimized inverted flat geometry (in Å). X is a label for the center of mass of the molecule.

1	C1	-15.015728	-1.1938718	-9.1716026
2	C2	-14.726166	-0.2530861	-10.1694960
3	C3	-14.028122	0.8266259	-9.6117327
4	C4	-13.216648	1.4309588	-7.3449705
5	C5	-13.014382	1.2561629	-5.9744097
6	C6	-13.408335	0.1933988	-5.1607788
7	C7	-14.103844	-0.9708566	-5.4892623
8	C8	-14.590518	-1.3781646	-6.7329156
9	C9	-14.503986	-0.7267595	-7.9545605
10	C10	-13.859995	0.5905520	-8.2414124
11	H1	-15.539195	-2.1168250	-9.3072465
12	H2	-14.997075	-0.3454082	-11.2013620
13	H3	-13.677984	1.6905151	-10.1364770
14	H4	-12.820593	2.3388077	-7.7634050
15	H5	-12.483516	2.0469373	-5.4789478
16	H6	-13.135365	0.2853835	-4.1256077
17	H7	-14.289666	-1.6449384	-4.6745772
18	H8	-15.101212	-2.3242710	-6.7475135
19	X	-14.039818	-0.0227078	-7.5707620

Listing S30: Atom coordinates for the optimized tilted geometry (in Å). X is a label for the center of mass of the molecule.

1	C1	-3.6573802	-11.1260710	13.1852123
2	C2	-3.9411660	-10.1823646	14.1820064
3	C3	-4.6333022	-9.0994265	13.6231375
4	C4	-5.4407573	-8.4926124	11.3556033
5	C5	-5.6422589	-8.6667935	9.9848515
6	C6	-5.2539342	-9.7324276	9.1722693
7	C7	-4.5670538	-10.9013218	9.5024309
8	C8	-4.0843632	-11.3109369	10.7468785
9	C9	-4.1666637	-10.6574668	11.9677127

10	C10	-4.8029223	-9.3360862	12.2531023
11	H1	-3.1391512	-12.0518305	13.3218420
12	H2	-3.6701514	-10.2747875	15.2138366
13	H3	-4.9780748	-8.2327633	14.1468590
14	H4	-5.8319284	-7.5822292	11.7731255
15	H5	-6.1674074	-7.8729032	9.4882806
16	H6	-5.5247077	-9.6391174	8.1366398
17	H7	-4.3850446	-11.5773077	8.6884630
18	H8	-3.5803263	-12.2605860	10.7627638
19	X	-4.6199504	-9.9475698	11.5816821

Listing S31: Atom coordinates for the optimized inverted tilted geometry (in Å). X is a label for the center of mass of the molecule.

1	C1	3.6573802	11.1260710	-13.1852123
2	C2	3.9411660	10.1823646	-14.1820064
3	C3	4.6333022	9.0994265	-13.6231375
4	C4	5.4407573	8.4926124	-11.3556033
5	C5	5.6422589	8.6667935	-9.9848515
6	C6	5.2539342	9.7324276	-9.1722693
7	C7	4.5670538	10.9013218	-9.5024309
8	C8	4.0843632	11.3109369	-10.7468785
9	C9	4.1666637	10.6574668	-11.9677127
10	C10	4.8029223	9.3360862	-12.2531023
11	H1	3.1391512	12.0518305	-13.3218420
12	H2	3.6701514	10.2747875	-15.2138366
13	H3	4.9780748	8.2327633	-14.1468590
14	H4	5.8319284	7.5822292	-11.7731255
15	H5	6.1674074	7.8729032	-9.4882806
16	H6	5.5247077	9.6391174	-8.1366398
17	H7	4.3850446	11.5773077	-8.6884630
18	H8	3.5803263	12.2605860	-10.7627638
19	X	4.6199504	9.9475698	-11.5816821

Table S5: Computed energetics for the azulene monomer. Oscillator strengths for the  $S_0 \rightarrow S_4$  and  $S_0 \rightarrow S_6$  transitions are not reported since below the  $10^{-5}$  threshold.

Azulene state	SA-CASSCF	CASPT2	Ref. weight	MS-CASPT2	Excitation energy		$f_{\text{Osc}}$	Einstein coefficient ( $s^{-1}$ )	TDM ( $e\text{\AA}$ ) (D)
	( $E_h, -383$ )	( $E_h, -384$ )	( $E_h, -384$ )	Main character	(eV)	(nm)			
$S_0$	-0.695 061	-1.156 753	0.68	-1.156 755	GS	—	—	—	—
$S_1$	-0.621 314	-1.077 262	0.68	-1.079 035	$\pi \rightarrow \pi^*(b_1)$	2.12	586	693 050	0.1389 0.667
$S_2$	-0.539 929	-0.972 257	0.68	-1.009 650	$\pi \rightarrow \pi^*(a_1)$	4.00	310	1 647 486	0.0822 0.395
$S_3$	-0.530 194	-1.009 652	0.67	-0.987 229	$\pi \rightarrow \pi^*(b_1)$	4.61	269	8 164 571	0.1480 0.711
$S_4$	-0.526 297	-0.956 504	0.68	-0.972 332	$\pi \rightarrow \text{Ryd}_s(a_2)$	5.02	247	—	0.0001 0.001
$S_5$	-0.524 235	-0.986 066	0.67	-0.956 504	$\pi \rightarrow \text{Ryd}_{p_y}(b_2)$	5.45	228	249 222	0.0201 0.097
$S_6$	-0.523 664	-0.954 636	0.68	-0.954 561	$\pi \rightarrow \text{Ryd}_{p_x}(a_2)$	5.50	225	—	0.0003 0.002
$S_7$	-0.520 226	-0.950 650	0.68	-0.947 715	$\pi \rightarrow \text{Ryd}_{p_z}(b_1)$	5.69	218	1 139 945	0.0404 0.194

### S9.1.2 Crystal spectrum evaluation

Since wave function orbital components are invariant to spatial rigid translation of the underlying atoms, we used the same four unique monomer MS-CASPT2 wave functions (i.e., the corresponding `JobMix` file) for all possible dimer evaluations. Out of the 500 azulene molecules contained in the  $5 \times 5 \times 5$  supercell, we chose one azulene monomer as reference. We located all symmetry unique dimers within 25 Å from the reference molecule (fig. S5), since we deemed that dimers further away would not have significant coupling terms. For these dimers, we computed excitonic couplings using the input shown in listing S34.

As a practical approximation when building the crystal excitonic Hamiltonian, we considered each monomer electronic excitation separately. This is justified by considering, for example, the flat–flat dimer that, among those we computed, has the largest computed excitonic couplings, in absolute value. As shown in table S6, excitonic states 2 and 3 are mostly constituted by a in-phase (state 2) and out-of-phase (state 3) combination of monomer excitations to the first excited state ( $\langle 1 2 |$  and  $\langle 2 1 |$ ). Since the in-phase one is the lowest excitation, these two states represent a  $J$ -aggregate, and the corresponding absorbance peak should be at least slightly red-shifted with respect to the monomer one. Similarly, excitonic states 6 and 7 are constituted by  $\langle 1 3 |$  and  $\langle 3 1 |$  (7 is in-phase, thus they are in a  $H$ -aggregate situation, blue-shift of the absorbance peak), and 8 and 9 are constituted by  $\langle 1 4 |$  and  $\langle 4 1 |$  (8 is in-phase, that is  $J$ -aggregate, red-shift of the absorbance peak). Contrarily, excitonic states 4, 5, 10, 11, 12, 13, 14, and 15 present a mixture of constituting states, also for those states with a clear predominance of a single monomer state (e.g., excitonic state 12). Taking into consideration the computed oscillator strengths for the excitonic excitation reported in table S7, only excitations to the excitonic states 2, 7 and 8 are worth further investigation. Therefore, the choice of using only the corresponding underlying monomer excitations, and not including couplings between different excitations.

Therefore, we built a 500 dimensional matrix, the diagonal elements of which were the monomer  $S_0 \rightarrow S_1$  excitation energy, and the off-diagonal elements were the computed Frenkel  $\langle (2)B(1)A | v | (1)B(2)A \rangle$  (i.e., monomer B is excited from state 1 to state 2, at the same time as monomer A is relaxed from state 2 to state 1) couplings between the various dimers. We employed periodic boundary conditions to introduce coupling between monomers at the cell unit border. Similarly, we built matrices using the monomer  $S_0 \rightarrow S_2$  excitation energy, along with  $\langle (3)B(1)A | v | (1)B(3)A \rangle$  couplings, and  $S_0 \rightarrow S_3$  excitation energy, along with  $\langle (4)B(1)A | v | (1)B(4)A \rangle$  couplings. After diagonalization of each matrix, the eigenvalues corresponded to the excitation energies, and applying the eigenvectors to the monomer transition dipole moments we obtained the corresponding oscillator strengths, used to plot the spectrum of fig. 18a.

Section S9.3 details how we used excitation energies and oscillator strengths data to plot the simulated spectra of azulene monomer and crystal.



Table S6: Eigenvectors of the Frenkel Hamiltonian relative to the flat-flat azulene dimer. Original values were in scientific notation, but are reported as float with only two decimals for ease of comparison. Yellow highlighted values are larger than 0.5 in absolute value, green highlighted values are comprised between 0.5 and 0.1, in absolute value. An excited excitonic state that was perfectly described by a combination of two equivalent monomer excited states (that is, as if in a dipole-dipole approximation) would have values  $\pm 1/\sqrt{2}$  for the combining states, and zeroes otherwise. An excitonic state that was mostly constituted by a single monomer excited state (that is, no exciton delocalization) would have values close to unity for such constituting state, and zeroes otherwise. For example, the excitonic ground state is completely constituted by the monomer ground states.

$\langle BA  $	Excitonic state														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\langle 11  $	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\langle 12  $	0.00	0.68	-0.73	0.03	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\langle 13  $	0.00	0.00	0.00	0.00	-0.01	-0.67	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\langle 14  $	0.00	0.00	0.00	0.11	-0.02	0.00	-0.01	0.72	-0.69	0.00	-0.02	0.00	0.00	0.00	0.00
$\langle 15  $	0.00	0.02	-0.02	-0.65	0.13	0.00	0.01	0.06	-0.06	0.01	0.72	0.12	0.01	0.14	-0.01
$\langle 16  $	0.00	-0.02	0.02	0.63	-0.12	0.00	-0.01	-0.04	0.04	0.01	0.64	0.18	0.01	-0.38	0.01
$\langle 17  $	0.00	0.00	0.00	-0.04	0.01	0.00	0.00	0.00	0.00	0.00	-0.21	0.98	0.05	0.04	0.00
$\langle 18  $	0.00	-0.01	0.01	0.36	-0.07	0.00	0.00	-0.02	0.02	0.00	0.16	0.01	0.00	0.91	-0.04
$\langle 21  $	0.00	0.73	0.68	0.01	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\langle 31  $	0.00	0.00	0.00	0.01	-0.02	0.75	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\langle 41  $	0.00	0.00	0.00	0.02	0.11	0.00	0.01	0.69	0.72	-0.02	0.00	0.00	0.00	0.00	0.00
$\langle 51  $	0.00	-0.02	-0.01	0.12	0.63	0.00	0.01	-0.06	-0.07	-0.74	0.01	-0.01	0.10	-0.01	-0.14
$\langle 61  $	0.00	-0.02	-0.02	0.12	0.63	0.00	0.01	-0.04	-0.04	0.59	-0.01	0.02	-0.31	-0.01	-0.37
$\langle 71  $	0.00	0.00	0.00	0.03	0.14	0.00	0.00	-0.01	-0.01	0.28	0.00	-0.05	0.94	-0.01	-0.11
$\langle 81  $	0.00	-0.01	-0.01	0.07	0.37	0.00	0.01	-0.02	-0.03	0.16	0.00	0.00	0.00	0.03	0.91

Table S7: Computed flat–flat azulene dimer spectral data, relative to only the excitations from the excitonic ground state.

from	to	exc. energy (eV)	osc. str.
1	2	2.11	0.0071
1	3	2.11	0.0000
1	4	3.69	0.0000
1	5	3.70	0.0001
1	6	4.00	0.0000
1	7	4.01	0.0047
1	8	4.62	0.0177
1	9	4.62	0.0000
1	10	5.16	0.0001
1	11	5.17	0.0002
1	12	5.36	0.0000
1	13	5.37	0.0000
1	14	5.67	0.0009
1	15	5.67	0.0006

## S9.2 Z–M azobenzene derivatives heterodimer

### S9.2.1 Monomer geometries and energetics

Geometries for the Z (fig. S6a) and M (fig. S6b) *trans*-azobenzene models were optimized in vacuo at the DFT level of theory using the B3LYP functional<sup>S17–S20</sup> and def2-TZVP basis set.<sup>S21,S22</sup> Both molecules presented a nearly perfectly planar geometry (fig. S6, panels c and d). Listings S32 and S33 report the optimized coordinates.

Single point energy evaluations were performed for both monomers at the four roots state-averaged RASSCF level of theory,<sup>S8</sup> using the double zeta ANO-RCC-VDZP basis set.<sup>S9</sup> The active orbitals comprised the whole conjugated  $\pi$  system, plus the two central nitrogen atoms’ lone pairs. Monomer Z comprised 5 orbitals in RAS1 (table S8), 7 in RAS2 (table S9), and 4 in RAS3 (table S10), with 3 holes/excitations allowed in RAS1 and RAS3, and a total of 18 active electrons. Monomer M comprised 6 orbitals in RAS1 (table S11), 7 in RAS2 (table S12), and 4 in RAS3 (table S13), with 3 holes and excitations allowed in RAS1 and RAS3, and a total of 20 active electrons. We used a level shift value of  $0.1 E_h$  to help wave function convergence. Subsequently, energies were re-evaluated at the MS-RASPT2 level,<sup>S23,S24</sup> with default IPEA shift value  $0.25 E_h$ ,<sup>S12</sup> and an imaginary level shift of  $0.1 E_h$ .<sup>S13</sup> Complete energetics data for Z and M azobenzene derivatives are reported in table S14, respectively.

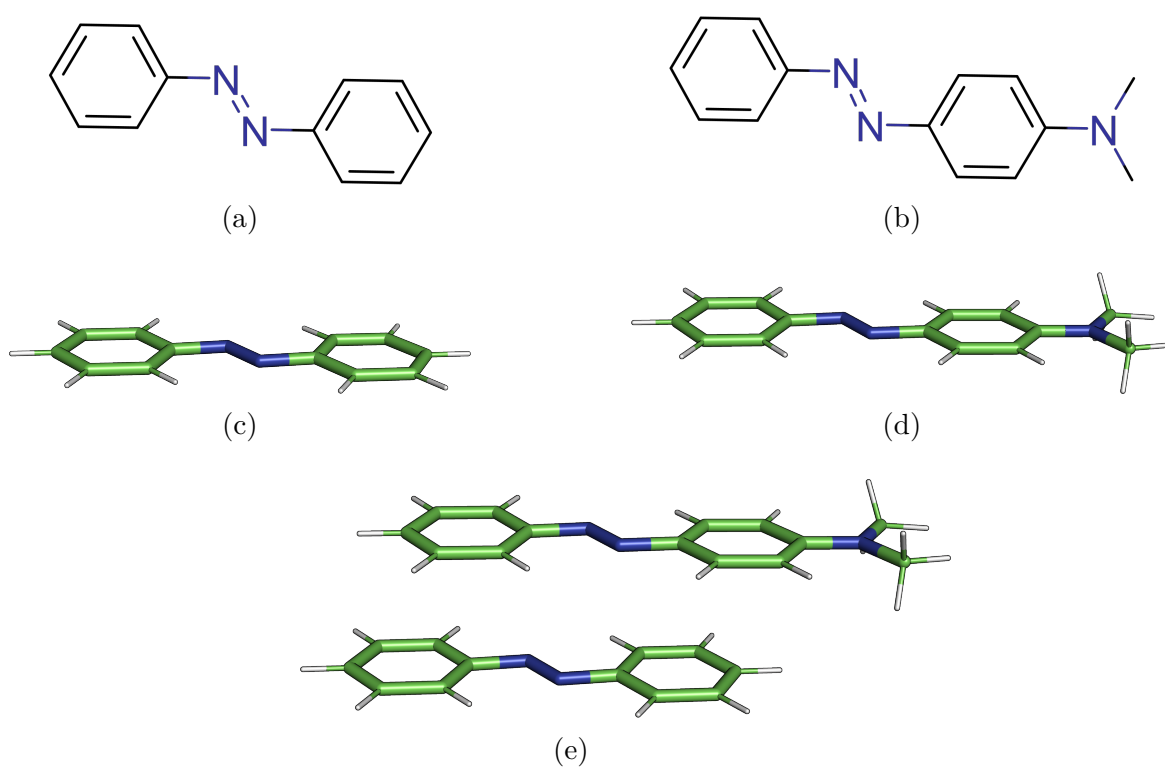


Figure S6: **Structures of *trans*-azobenzene models.** 2D chemical structures of Z (a) and M (b) compounds, while (c) and (d) are 3D representations of the corresponding final optimized geometries. ZM dimer model is shown in (e) where Z and M are in antiparallel orientation and separated by a distance of 3.5 Å.

Listing S32: Coordinates (in Å) of Z final optimized geometry. The coordinates of the centre of mass are (5.898, -0.350, -1.271).

1	C	5.89790438	-0.40458092	3.25861401
2	C	5.89789351	0.86664170	2.69687206
3	C	5.89791048	1.01261834	1.31596003
4	C	5.89785772	-1.53073770	2.43471036
5	C	5.89788167	-1.39274702	1.05718685
6	C	5.89789107	-0.11241877	0.49031776
7	N	5.89789467	0.15269137	-0.90112465
8	N	5.89789536	-0.85211074	-1.64086575
9	C	5.89789381	-0.58703052	-3.03231030
10	C	5.89787763	0.69327808	-3.59922482
11	C	5.89785058	0.83121533	-4.97675526
12	C	5.89790632	-1.71209810	-3.85791172
13	C	5.89790731	-1.56617704	-5.23882809
14	C	5.89789921	-0.29497443	-5.80061572
15	H	5.89790728	1.74240821	3.33314577
16	H	5.89793014	1.99098590	0.85336050
17	H	5.89786037	-2.52002729	2.87531099
18	H	5.89787024	-2.25586886	0.40709428
19	H	5.89785977	1.55642495	-2.94916460
20	H	5.89787929	1.82048877	-5.41739498
21	H	5.89787168	-0.17780331	-6.87697021
22	H	5.89792363	-2.69044626	-3.39527172
23	H	5.89791186	-2.44196968	-5.87506823
24	H	5.89788201	-0.52180199	4.33496345

Listing S33: Coordinates (in Å) of M final optimized geometry. The coordinates of the centre of mass are (9.1922, -0.49978, 0.03962).

1	C	9.16908905	-0.64135276	-4.47233796
2	C	9.13228658	0.51481030	-3.69264092
3	C	9.14583343	0.43250918	-2.31039729
4	C	9.21830312	-1.88723733	-3.85774088
5	C	9.23053558	-1.97715663	-2.47201689
6	C	9.19599677	-0.82181523	-1.68873581
7	N	9.21038293	-1.02779639	-0.28864668
8	N	9.21221652	0.01585924	0.40478472
9	C	9.21924663	-0.16350371	1.79417600
10	C	9.23639225	0.99853690	2.56958559
11	C	9.24268980	0.94596169	3.94950316
12	C	9.20747095	-1.40147673	2.45337661
13	C	9.21175945	-1.46848733	3.82881943
14	C	9.22959523	-0.29438819	4.62313133
15	H	9.09202143	1.48565937	-4.17139596
16	H	9.11708573	1.32115177	-1.69632107
17	H	9.24636804	-2.78843665	-4.45740342
18	H	9.26730280	-2.93624160	-1.97163698
19	H	9.24484717	1.95370770	2.06004376
20	H	9.25591933	1.87015957	4.50698743
21	N	9.23305729	-0.36293613	5.99375562
22	H	9.19370851	-2.30763180	1.86392392
23	H	9.20111860	-2.44045913	4.29926610
24	C	9.20556873	-1.64837593	6.66305485
25	C	9.25558534	0.85058773	6.78644498
26	H	10.07688706	-2.25861216	6.40823609
27	H	8.30929363	-2.22047592	6.40624120
28	H	9.20792090	-1.49295118	7.73869511

29	H	10.14925763	1.45069763	6.58552188
30	H	9.25859114	0.58767217	7.84109055
31	H	8.37781336	1.47756579	6.59718525
32	H	9.15817501	-0.56720423	-5.55251971

Table S8: RAS1 active space orbitals for *trans*-azobenzene Z monomer. Electronic occupation for each state is given under each orbital. Density contour value = 0.05.


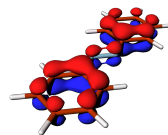
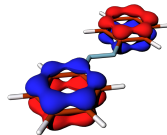
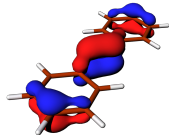
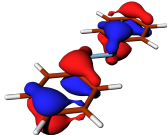
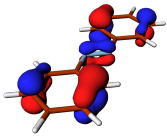
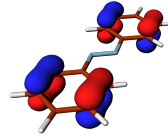
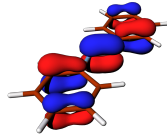
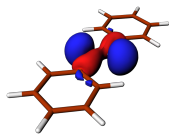
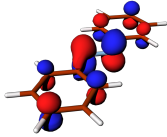
					
S <sub>0</sub>	1.998	1.965	1.965	1.929	1.912
S <sub>1</sub>	1.944	1.968	1.966	1.949	1.919
S <sub>2</sub>	1.998	1.941	1.927	1.904	1.791
S <sub>3</sub>	1.998	1.940	1.927	1.904	1.788

Table S9: RAS2 active space orbitals for *trans*-azobenzene Z monomer. Electronic occupation for each state is given under each orbital. Density contour value = 0.05.

					
S <sub>0</sub>	1.900	1.901	1.854	1.997	0.153
S <sub>1</sub>	1.901	1.901	1.884	1.055	1.051
S <sub>2</sub>	1.620	1.747	1.506	1.997	0.580
S <sub>3</sub>	1.745	1.520	1.611	1.997	0.577



		
S <sub>0</sub>	0.100	0.100
S <sub>1</sub>	0.108	0.101
S <sub>2</sub>	0.330	0.242
S <sub>3</sub>	0.244	0.328

Table S10: RAS3 active space orbitals for *trans*-azobenzene Z monomer. Electronic occupation for each state is given under each orbital. Density contour value = 0.05.

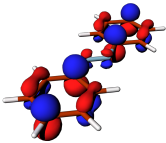
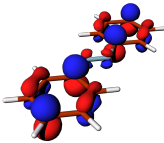
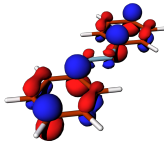
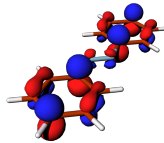
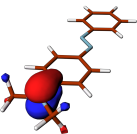
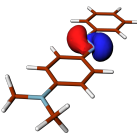
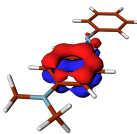
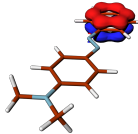
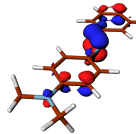
				
S <sub>0</sub>	0.090	0.072	0.033	0.032
S <sub>1</sub>	0.099	0.091	0.034	0.034
S <sub>2</sub>	0.217	0.091	0.059	0.050
S <sub>3</sub>	0.220	0.091	0.059	0.050

Table S11: RAS1 active space orbitals for *trans*-azobenzene M monomer. Electronic occupation for each state is given under each orbital. Density contour value = 0.05.

					
S <sub>0</sub>	1.997	1.998	1.967	1.965	1.932
S <sub>1</sub>	1.998	1.944	1.969	1.967	1.950
S <sub>2</sub>	1.994	1.998	1.934	1.965	1.907
S <sub>3</sub>	1.997	1.998	1.967	1.934	1.904

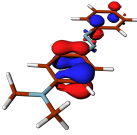
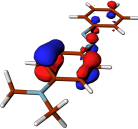
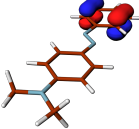
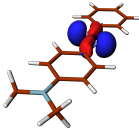
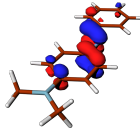
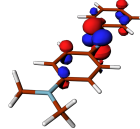
	
S <sub>0</sub>	1.915
S <sub>1</sub>	1.922
S <sub>2</sub>	1.857
S <sub>3</sub>	1.860

Table S12: RAS2 active space orbitals for *trans*-azobenzene M monomer. Electronic occupation for each state is given under each orbital. Density contour value = 0.05.

					
S <sub>0</sub>	1.909	1.909	1.997	1.853	0.159
S <sub>1</sub>	1.909	1.909	1.055	1.884	1.049
S <sub>2</sub>	1.581	1.581	1.997	1.328	0.460
S <sub>3</sub>	1.901	1.901	1.997	1.499	0.669

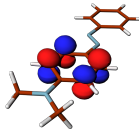
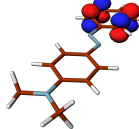
		
S <sub>0</sub>	0.093	0.100
S <sub>1</sub>	0.097	0.101
S <sub>2</sub>	0.678	0.108
S <sub>3</sub>	0.107	0.474

Table S13: RAS3 active space orbitals for *trans*-azobenzene M monomer. Electronic occupation for each state is given under each orbital. Density contour value = 0.05.

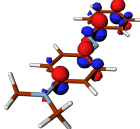
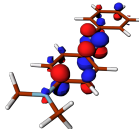
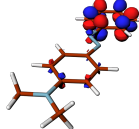
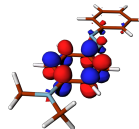
				
S <sub>0</sub>	0.087	0.069	0.033	0.031
S <sub>1</sub>	0.098	0.090	0.034	0.032
S <sub>2</sub>	0.126	0.088	0.032	0.063
S <sub>3</sub>	0.130	0.085	0.063	0.031

Table S14: Complete energetics data and properties for azobenzene Z and M derivatives.

Z monomer	SA-CASSCF	RASPT2	Ref. weight	MS-RASPT2	Excitation energy	$f_{\text{Osc}}$	Einstein coefficient	TDM
state	$(E_h, -569)$	$(E_h, -571)$		$(E_h, -571)$	Main character (eV) (nm)		$(\text{s}^{-1})$	$(e\text{\AA})$ (D)
S <sub>0</sub>	-0.683515	-0.432985	0.63	-0.432985	GS	—	—	—
S <sub>1</sub>	-0.563434	-0.337809	0.62	-0.337809	$n \rightarrow \pi^*$	2.59 479	0.00000	—
S <sub>2</sub>	-0.506697	-0.267140	0.62	-0.267140	$\pi \rightarrow \pi^*$	4.51 275	0.01204	10642888 0.17 0.84
S <sub>3</sub>	-0.506165	-0.266114	0.62	-0.266113	$\pi \rightarrow \pi^{*2}$	4.54 273	0.00000	—

M monomer	SA-CASSCF	RASPT2	Ref. weight	MS-RASPT2	Excitation energy	$f_{\text{Osc}}$	Einstein coefficient	TDM
state	$(E_h, -702)$	$(E_h, -704)$		$(E_h, -704)$	Main character (eV) (nm)		$(\text{s}^{-1})$	$(e\text{\AA})$ (D)
S <sub>0</sub>	-0.859348	-1.089153	0.57	-1.089223	GS	—	—	—
S <sub>1</sub>	-0.738453	-0.991638	0.56	-0.991638	$n \rightarrow \pi^*$	2.66 467	0.00005	14488 0.01 0.07
S <sub>2</sub>	-0.685583	-0.931741	0.56	-0.931700	$\pi \rightarrow \pi^*$	4.29 289	0.03550	28301885 0.31 1.48
S <sub>3</sub>	-0.682774	-0.923560	0.56	-0.923530	$\pi \rightarrow \pi^{*2}$	4.51 275	0.00538	4742679 0.12 0.56



### S9.2.2 Excitonic spectrum evaluation

A ZM heterodimer model (fig. S6e) was constructed as close as possible to what described in ref. S25: Z and M were parallel, with the respective center of masses at a distance of 3.5 Å. Excitonic coupling was computed using the input reported as listing S35.

Obtained Frenkel’s Hamiltonian eigenvectors, and corresponding excitonic excitation energies are reported in tables S15 and S16. As opposed to the data presented in table S6, ZM excitonic states 2, 3, 4, and 5 are dominated by a single state. In particular, excitonic state 2 is dominated by the second excited state of the Z monomer, while states 3, 4, and 5 are dominated by the second, third, and fourth excited states of the M monomer, respectively. Finally, excitonic states 6 and 7 are due to a mixing of M monomer third and fourth excited states, which are rather close in energy.

The data was collated to plot the spectrum, following section S9.3, and to draw the state diagram shown in fig. S7. It is notable how all excitonic excited states are red-shifted with respect to the originating monomer states. A dipole–dipole approximation based treatment of the system would not be capable of describing this peculiar situation, since it would only couple states of different monomers to induce splitting.

Table S15: Eigenvectors of the Frenkel Hamiltonian relative to ZM heterodimer. Original values were in scientific notation, but are reported as float with only two decimals for ease of comparison. Yellow highlighted values are larger than 0.5 in absolute value, green highlighted values are comprised between 0.5 and 0.1, in absolute value. *A* stands for the states of the Z monomer, *B* stands for the states of the M monomer.

$\langle BA $	Excitonic state						
	1	2	3	4	5	6	7
$\langle 11 $	1.00	0.01	0.00	-0.01	0.02	0.02	0.00
$\langle 12 $	-0.01	0.99	-0.13	0.00	0.00	0.00	0.00
$\langle 13 $	-0.02	0.00	0.00	-0.02	-0.01	0.91	0.42
$\langle 14 $	0.00	0.00	0.00	0.01	0.02	-0.42	0.91
$\langle 21 $	0.00	0.13	0.99	0.00	0.00	0.00	0.00
$\langle 31 $	0.01	0.00	0.00	1.00	0.00	0.02	0.00
$\langle 41 $	-0.02	0.00	0.00	0.00	1.00	0.02	-0.01

Table S16: ZM dimer spectral data, relative to only the excitations from the excitonic ground state.

from	to	exc. energy (eV)	osc. str.
1	2	2.49	0.0000
1	3	2.56	0.0000
1	4	3.85	0.0313
1	5	4.01	0.0049
1	6	4.04	0.0087
1	7	4.07	0.0021

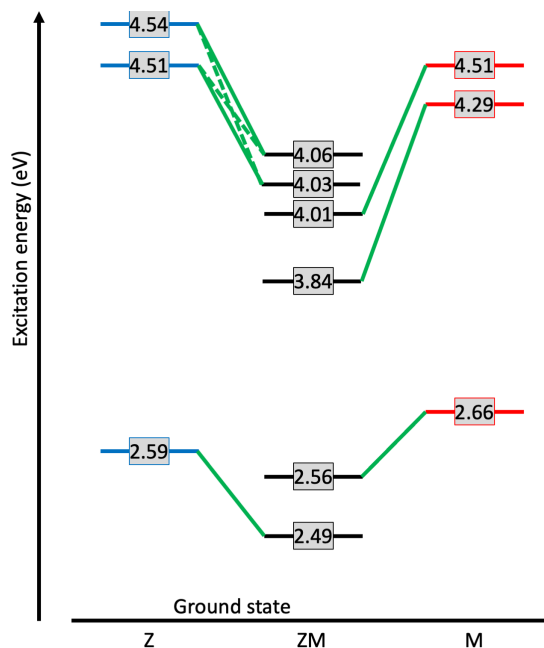


Figure S7: **State diagram for ZM heterodimer.** Z monomer excited states (blue boxes on the left) and M monomer excited states (red boxes on the right) are combined into the ZM heterodimer excitonic states (black boxes in the middle). Excitation energies are on a qualitative scale. All heterodimer excitonic excited states are red-shifted with respect to the generating monomer states.

### S9.3 Plotting absorption spectra

Starting from a set of excitation energies and their corresponding oscillator strengths  $f$ , we plotted a simulated absorption spectrum as here detailed. We generated a Gaussian-like distribution for each excitation, with the line shape function defined as:

$$g(E) = \frac{f_n}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(E - \Delta E_{S_n-S_0})^2}{2\sigma^2}\right) \quad (\text{S3})$$

Each function is centered at the excitation value ( $\Delta E_{S_n-S_0}$ ) and weighted by the corresponding  $f_n$ , with  $n$  running over all excitations of the considered molecule/system. The parameter  $\sigma$  is a phenomenological bandwidth, which we chose to be 0.1 eV. The final spectrum results from the sum of Gaussian-like distributions.

In fig. 18a, the azulene monomer spectrum was plotted using the data of table S5, while the one relative to the crystal used what obtained through the procedure described at the end of section S9.1.2. Both spectra were normalized such as the highest peak of the monomer spectrum corresponded to an intensity of 1.

In fig. 18b, the absorbance spectrum for Z and M monomers was plotted using the data of table S14, and for ZM heterodimer the data from table S16. All spectra were normalized such that the M monomer highest peak intensity corresponded to 1.

### S9.4 Example input files

We present here the input files required to compute Frenkel's excitonic couplings for a homodimer of azulene (listing S34) and the heterodimer of Z and M *trans*-azobenzene (listing S35). In both cases, energetics calculations for each monomer have already been

performed. Further technical details on the use of the RUNFILE, as well as when to use the keywords *ricd* and *fake ricd* are given in ref. S26.

Listing S34: Example of input for a dimer of azulene molecules. The input implies that one has already computed the energetics for both monomers, and has available the corresponding geometries and JobMix files.

```
1 > COPY $InputDir/flat_opt_248.xyz monoA.xyz
2 > COPY $InputDir/tilt_opt_231.xyz monoB.xyz
3
4 ***** INTEGRALS OF A+B PART 1 *****
5
6 &GATEWAY
7 BASDIR = $InputDir
8 coord=monoA.xyz
9 coord=monoB.xyz
10 bsse=1
11 Basis=
12 ANO-RCC-VTZP, X.AZURYDBERG
13 NoMove
14 Group=Nosym
15 ricd
16
17 &SEWARD
18
19 >>> COPY $Project.RunFile AUXRFIL1
20
21 ***** SYSTEM A *****
22
23 &GATEWAY
24 BASDIR = $InputDir
25 coord=monoA.xyz
26 Basis=
27 ANO-RCC-VTZP, X.AZURYDBERG
28 NoMove
29 Group=Nosym
30
31 &SEWARD
32 fake ricd
33
34 > COPY $InputDir/flat_opt-pt2_mix.JobMix JOB001
35
36 &RASSI
37 EJOB
38 Nr of Job = 1 all
39 TRD1
40 TRDI
41 MONA
42
43 ***** INTEGRALS OF A+B PART 2 *****
44
45 &GATEWAY
46 BASDIR = $InputDir
47 coord=monoA.xyz
48 coord=monoB.xyz
49 bsse=2
50 Basis=
51 ANO-RCC-VTZP, X.AZURYDBERG
```

```

52 NoMove
53 Group=Nosym
54
55 &SEWARD
56 fake ricd
57
58 >>> COPY $Project.RunFile AUXRFIL2
59
60 ***** SYSTEM B *****
61
62 &GATEWAY
63 BASDIR = $InputDir
64 coord=monoB.xyz
65 Basis=
66 ANO-RCC-VTZP, X.AZURYDBERG
67 NoMove
68 Group=Nosym
69
70 &SEWARD
71 fake ricd
72
73 > COPY $InputDir/tilt_opt-pt2_mix.JobMix JOB001
74
75 &RASSI
76 EJOB
77 Nr of Job = 1 all
78 TRD1
79 TRDI
80 MONB
81 EXCItionics

```

Listing S35: Example of input for a heterodimer of Z and M *trans*-azobenzene molecules. The input implies that one has already computed the energetics for both monomers, and has available the corresponding geometries and JobMix files.

```

1 > COPY $InputDir/monoZ.xyz monoA.xyz
2 > COPY $InputDir/monoM.xyz monoB.xyz
3
4 ***** INTEGRALS OF A+B PART 1 *****
5
6 &GATEWAY
7 coord=monoA.xyz
8 coord=monoB.xyz
9 bsse=1
10 Basis=ANO-RCC-VDZP
11 NoMove
12 Group=Nosym
13 ricd
14
15 &SEWARD
16
17 >>> COPY $Project.RunFile AUXRFIL1
18
19 ***** SYSTEM A *****
20
21 &GATEWAY
22 coord=monoA.xyz
23 Basis=ANO-RCC-VDZP

```

```

24 NoMove
25 Group=Nosym
26
27 &SEWARD
28 fake ricd
29
30 > COPY $InputDir/CPL.monoZ.JobMix JOB001
31
32 &RASSI
33 EJOB
34 Nr of Job = 1 all
35 TRD1
36 TRDI
37 MONA
38
39 ***** INTEGRALS OF A+B PART 2 *****
40
41 &GATEWAY
42 coord=monoA.xyz
43 coord=monoB.xyz
44 bsse=2
45 Basis=ANO-RCC-VDZP
46 NoMove
47 Group=Nosym
48
49 &SEWARD
50 fake ricd
51
52 >>> COPY $Project.RunFile AUXRFIL2
53
54 ***** SYSTEM B *****
55
56 &GATEWAY
57 coord=monoB.xyz
58 Basis=ANO-RCC-VDZP
59 NoMove
60 Group=Nosym
61
62 &SEWARD
63 fake ricd
64
65 > COPY $InputDir/CPL.MonoM.JobMix JOB001
66
67 &RASSI
68 EJOB
69 Nr of Job = 1 all
70 TRD1
71 TRDI
72 MONB
73 EXCItonics

```

## S10 Ionization and Autoionization Processes

This section reports examples of calculations of single-site double-core-hole (ssDCH) and Auger decay spectra. The implementation details can be found in refs. S27, S28, respec-

tively.

## S10.1 ssDCH

For the calculation of double-core-hole spectral intensities, target double-core-hole state wave functions are computed using the *DEXS* keyword on the *RASSCF* input. Below we give an input example for the calculation of the ssDCH spectrum of CO at the carbon  $K^{-2}$  edge for 30 final ssDCH states. Notice that our approach is invoked with the keyword *DCHS* in the *RASSI* input section, followed by an integer that specifies the orbital number of the double core hole (normally this should be equal to 1, that is, the first active orbital).

Listing S36: Input example of a ssDCH spectrum calculation.

```
1  &GATEWAY
2  expert
3  SYMMetry
4  xy y
5  Basis set
6  0.AUG-CC-PVQZ
7  Spherical all
8  0                0.0000000000        0.0000000000        -1.0658055343 / Bohr
9  End of basis
10
11 Basis set
12 C.AUG-CC-PVQZ..
13 Spherical all
14 C                0.0000000000        0.0000000000        1.0658055343 / Bohr
15 End of Basis
16 Verbose
17 BSShow
18
19 &SEWARD
20 Cholesky
21
22 &SCF
23
24 &RASSCF &END
25 EXPERT
26 LUMORB
27 Symmetry = 1
28 Spin = 1
29 nActEl = 10 2 0
30 ALTER
31 1
32 1 2 3
33 SUPSym
34 1
35 1 2
36 0; 0; 0
37 Inactive = 2 0 0 0
38 Ras1 = 1 0 0 0
39 Ras2 = 4 3 0 3
40 Ras3 = 0 0 0 0
41 LEVShft = 0.7
42 CIR0ot = 1 1 1
43 Iter=200 100
44 End of input
```

```

45 >> COPY $Project.JobIph $Project.JobIph_GS
46
47 >> COPY $Project.RasOrb INPORB
48 &RASSCF &END
49 EXPERT
50 LUMORB
51 Symmetry = 1
52 Spin = 1
53 nActEl = 8 2 0
54 SUPSym
55 1
56 1 3
57 0; 0; 0
58 DEXS
59 1
60 1
61 Inactive = 2 0 0 0
62 Ras1 = 1 0 0 0
63 Ras2 = 4 3 0 3
64 Ras3 = 0 0 0 0
65 LEVShft = 1.2
66 OrbListing = All
67 OrbAppear = Full
68 Ciroot = 30 30 1
69 >> COPY $Project.JobIph $Project.JobIph_A1x
70
71 >> COPY $Project.JobIph_GS JOBIPH
72 &CASPT2 &END
73 MAXITER
74 60
75 MULTistate = all
76 IMAG
77 0.25
78 End of input
79 >> COPY $Project.JobMix JOB001
80
81 >> COPY $Project.JobIph_A1x JOBIPH
82 &CASPT2 &END
83 MAXITER
84 60
85 MULTistate = all
86 IMAG
87 0.25
88 End of input
89 >> COPY $Project.JobMix JOB002
90
91 &RASSI
92 Nrof JobIphs = 2 all
93 Ejob
94 DCHS = 1
95 End of input

```

The output of the RASSI section shows the binding energies and spectral intensities, which are used to plot the ssDCH spectrum of fig. S8.

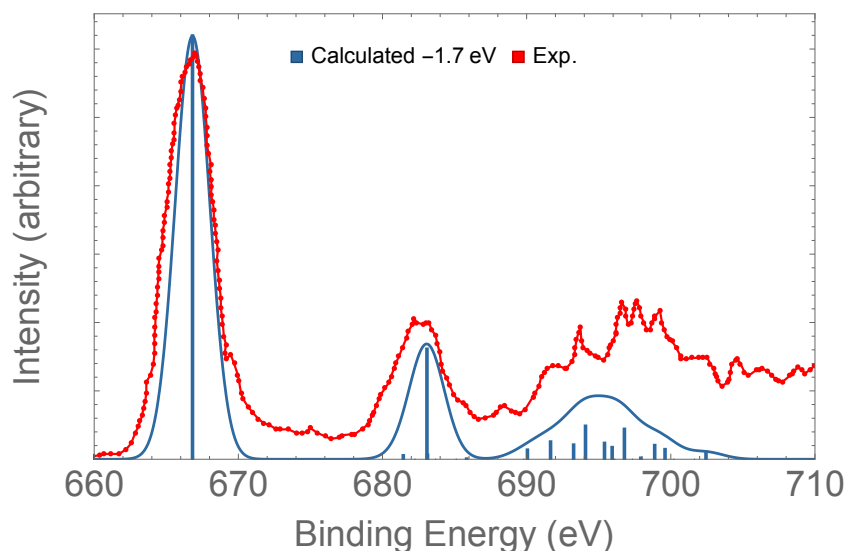


Figure S8: ssDCH stick spectrum of CO at the carbon  $K^{-2}$  edge, convoluted with Gaussian functions using half-width-at-half-maximum (hwhm) of 0.45 eV. The binding energies have been shifted by  $-1.3$  eV to better compare to the experiment. The experimental points were extracted from ref. S29.

## S10.2 Auger Decay Spectra

Auger spectra are obtained within the one-center approximation<sup>S28</sup> (OCA), which is currently only available for K-edge calculations of second and third row elements (C, N, O, F, Ne, Na, Mg, Al, Si, P, S, Cl, Ar). Auger matrix elements are computed with the *TDYS* keyword in the *RASSI* input section, followed by an integer ( $n$ ) specifying the number of scattering centers. The next  $n$  lines contain strings with the scattering center type. An example for the computation of Auger matrix elements for the carbon K-edge is: *TDYS= 1; C 1s*. Notice that *TDYS* requires the *DYSON* keyword and support to HDF5 files.

The next step is to run the *AugerOca* program, which can be found in the *Tools* folder. *AugerOca* requires Python 3 and support to *numpy* and *h5py* packages. The *AugerOca* Python 3 package (as well as *numpy* or *h5py*) can be installed via *pip*.

```
1 $pip install auger-oca
```

After installing the required packages, the python script *auger\_main.py* can be copied and executed from any directory. For basic instructions and options, see the help menu.

### Listing S37: Help menu of *auger\_main.py*.

```
1 $python3 auger_main.py -h
2
3 usage: auger_main.py [-h] [-i INP] [-d DIRECTORY] [--oca] [--spec] [--raes] [--aes] [--t]
4 ↪ [-s]
5
6 This program will process the 2p-Dyson density files (r2TM_) and provide:
7 Auger decay rates via one-center approach (OCA).
8 [Remember to make sure you have the $Project.rassi.h5 file in your directory.]
9
10 usage example (RAES):
11
12 $python3 $Tools/AugerOca/auger_main.py --raes -i r2TM_K2V_002_001
13
14 usage example (AES for a triplet final state):
```



```

14
15 $python3 $Tools/AugerOca/auger_main.py --aes --t -i r2TM_K2V_002_001
16
17 (recommended) usage based on a directory with a collection of r2TM_ files.
18 This directory can be the scratch directory for the RASSI calculation.
19
20 $python3 $MOLCAS/Tools/AugerOca/auger_main.py --raes --spec -d $WorkDir
21
22 options:
23 -h, --help            show this help message and exit
24 -i INP, --input INP  path of the density file (r2TM_)
25 -d DIRECTORY, --directory DIRECTORY
26                     path of a directory with a number of density files (r2TM_)
27 --oca                Perform Auger OCA (default: True)
28 --spec              save auger.spectrum.out (default: False)
29 --raes              Consider resonant Auger [RAES] (default: True if OCA)
30 --aes               Consider non-resonant Auger [AES] (default: False)
31 --t                 For AES, consider triplet final states (default: False)
32 --s                 For AES, consider singlet final states (default: False; True if AES)

```

Below we give one input example to compute the resonant Auger spectrum (RAES) of the  $N_2$  molecule. Here, the AugerOca post processing step is incorporated in the input of the calculation with a  $\gg$ *SHELL* command.

For practical reasons, for molecules with equivalent scattering centers we recommend to localize the core orbitals, preferably using a Cholesky localization procedure (maintaining symmetry), or any other localization procedure such as Boys or Pipek–Mezey (lifting symmetry point group). See ref. S28 for a detailed description of the RAES spectrum calculation of  $N_2$ .

Listing S38: Input example of a resonant Auger spectrum calculation. Notice that the input file shown is incomplete. It must be completed following the \* *REPEAT* instructions.

```

1  *-----*
2  * Molecule: N2
3  * Basis: AUG-CC-PVQZ
4  * Symmetry: C2v with localized core
5  * Features tested: N2 resonant Auger spectrum using auger_oca package
6  * Comments: auger_oca package, if not available, can be installed with your prompt:
7  * pip install auger_oca
8  * N 1s orbitals are localized with the Cholesky method
9  *-----*
10
11  &GATEWAY
12  expert
13  SYMMetry
14  xy y
15  Basis set
16  N.aug-cc-pvqz..
17  Spherical all
18  n1          0.0000000000    0.0000000000    0.548800 / Angstrom
19  End of basis
20
21  Basis set
22  N.aug-cc-pvqz..
23  Spherical all
24  n2          0.0000000000    0.0000000000   -0.548800 / Angstrom
25  End of basis
26  Verbose
27  BSShow
28

```

```

29  &SEWARD &END
30  Cholesky
31  End of input
32
33  &SCF &END
34  End of input
35
36  &LOCALISATION
37  NORBitals
38  2 0 0 0
39  Cholesky
40
41  &RASSCF
42  EXPERT
43  LUMORB
44  Symmetry
45  1
46  Spin
47  1
48  nActEl
49  14 1 0
50  SUPSym
51  2
52  1 2
53  1 1
54  0
55  0
56  0
57  Inactive
58  0 0 0 0
59  Ras1
60  2 0 0 0
61  Ras2
62  4 2 0 2
63  Ras3
64  0 0 0 0
65  CIRO
66  1 1 1
67  Iter=200 100
68  End of input
69  >>> COPY $Project.JobIph $Project.JobIph_gs
70  >>> COPY $Project.RasOrb $Project.RasOrb_gs
71
72  >> COPY $Project.RasOrb_gs INPORB
73  &RASSCF &END
74  EXPERT
75  LUMORB
76  Symmetry
77  4
78  Spin
79  1
80  nActEl
81  14 1 0
82  SUPSym
83  2
84  1 1
85  1 2
86  0

```

```

87 0
88 0
89 HEXS
90 1
91 1
92 Inactive
93 0 0 0 0
94 Ras1
95 2 0 0 0
96 Ras2
97 4 2 0 2
98 Ras3
99 0 0 0 0
100 CIRO
101 6 6 1
102 LEVShft
103 0.8
104 Iter=200 100
105 End of input
106 >> COPY $Project.JobIph $Project.JobIph_CORE
107
108 >> COPY $Project.RasOrb_gs INPORB
109 &RASSCF &END
110 EXPERT
111 LUMORB
112 Symmetry
113 1
114 Spin
115 2
116 nActEl
117 13 1 0
118 SUPSym
119 2
120 1 1
121 1 2
122 0
123 0
124 0
125 Inactive
126 0 0 0 0
127 Ras1
128 2 0 0 0
129 Ras2
130 4 2 0 2
131 Ras3
132 0 0 0 0
133 CIRO
134 30 30 1
135 Iter=200 100
136 End of input
137 >> COPY $Project.JobIph $Project.JobIph_A1
138
139 >> COPY $Project.RasOrb_gs INPORB
140 &RASSCF &END
141 EXPERT
142 LUMORB
143 Symmetry
144 2

```

```

145 Spin
146 2
147 nActEl
148 13 1 0
149 SUPSym
150 2
151 1 1
152 1 2
153 0
154 0
155 0
156 Inactive
157 0 0 0 0
158 Ras1
159 2 0 0 0
160 Ras2
161 4 2 0 2
162 Ras3
163 0 0 0 0
164 CIRO
165 30 30 1
166 Iter=200 100
167 End of input
168 >> COPY $Project.JobIph $Project.JobIph_B1
169
170 >> COPY $Project.RasOrb_gs INPORB
171 &RASSCF &END
172 EXPERT
173 LUMORB
174 Symmetry
175 4
176 Spin
177 2
178 nActEl
179 13 1 0
180 SUPSym
181 2
182 1 1
183 1 2
184 0
185 0
186 0
187 Inactive
188 0 0 0 0
189 Ras1
190 2 0 0 0
191 Ras2
192 4 2 0 2
193 Ras3
194 0 0 0 0
195 CIRO
196 30 30 1
197 Iter=200 100
198 End of input
199 >> COPY $Project.JobIph $Project.JobIph_B2
200
201 >> COPY $Project.RasOrb_gs INPORB
202 &RASSCF &END

```

```

203 EXPERT
204 LUMORB
205 Symmetry
206 3
207 Spin
208 2
209 nActEl
210 13 1 0
211 SUPSym
212 2
213 1 1
214 1 2
215 0
216 0
217 0
218 Inactive
219 0 0 0 0
220 Ras1
221 2 0 0 0
222 Ras2
223 4 2 0 2
224 Ras3
225 0 0 0 0
226 CIRO
227 30 30 1
228 Iter=200 100
229 End of input
230 >> COPY $Project.JobIph $Project.JobIph_A2
231
232
233 >> COPY $Project.JobIph_CORE JOBIPH
234 &CASPT2 &END
235 MAXITER
236 160
237 MULTistate = 1 1
238 IMAG
239 0.25
240 End of input
241 >> COPY $Project.JobMix $Project.JobMix_CORE
242 >> COPY $Project.JobMix_CORE JOB001
243
244 >> COPY $Project.JobIph_A1 JOBIPH
245 &CASPT2 &END
246 MAXITER
247 160
248 MULTistate = all
249 IMAG
250 0.25
251 End of input
252 >> COPY $Project.JobMix $Project.JobMix_A1x
253 >> COPY $Project.JobMix_A1x JOB002
254
255 >> COPY $Project.JobIph_B1 JOBIPH
256 &CASPT2 &END
257 MAXITER
258 160
259 MULTistate = all
260 IMAG

```

```

261 0.25
262 End of input
263 >> COPY $Project.JobMix $Project.JobMix_B1x
264 >> COPY $Project.JobMix_B1x JOB003
265
266 >> COPY $Project.JobIph_B2 JOBIPH
267 &CASPT2 &END
268 MAXITER
269 160
270 MULTistate = all
271 IMAG
272 0.25
273 End of input
274 >> COPY $Project.JobMix $Project.JobMix_B2x
275 >> COPY $Project.JobMix_B2x JOB004
276
277 >> COPY $Project.JobIph_A2 JOBIPH
278 &CASPT2 &END
279 MAXITER
280 160
281 MULTistate = all
282 IMAG
283 0.25
284 End of input
285 >> COPY $Project.JobMix $Project.JobMix_A2x
286 >> COPY $Project.JobMix_A2x JOB005
287
288 &RASSI
289 Nrof JobIphs = 5 all
290 Ejob
291 TDYS
292 2
293 n1 1s
294 n2 1s
295 DYSON
296 End of input
297
298 >> SHELL $MOLCAS/Tools/Auger0ca/auger_main.py -d $WorkDir --raes --spec
299 >> COPY auger.spectrum.out $CurrDir/$Project.auger.spectrum.out

```

After the Happy Landing, a file `auger.spectrum.out` containing the Auger spectrum is created. The data is used to plot the spectrum in fig. S9.

Listing S39: First five lines of the `auger.spectrum.out` file.

```

1 385.144525272 4.703085550629902e-05 4.806396806165732e-05
2 382.040239436 1.4554513122060067e-06 1.4858744308002272e-06
3 375.405751766 5.773528972561873e-05 5.9010026834452716e-05
4 374.625644331 3.840425866689331e-05 3.924869159969684e-05
5 374.402025037 6.288330829760877e-06 6.42586976974955e-06
6 ...

```

Notice that the first column of the `auger.spectrum.out` file gives the kinetic energy, followed by the Auger intensity on each nitrogen atom. As the nitrogen atoms are equivalent in this example, the intensities on the same row are expected to be similar.

We recommend using the built-in basis sets available within the `OpenMolcas` package, especially the Dunning and ANO families of basis sets. Using tailor-made basis sets or basis sets imported from external repositories, such as `Basis Set Exchange (BSE)`, is

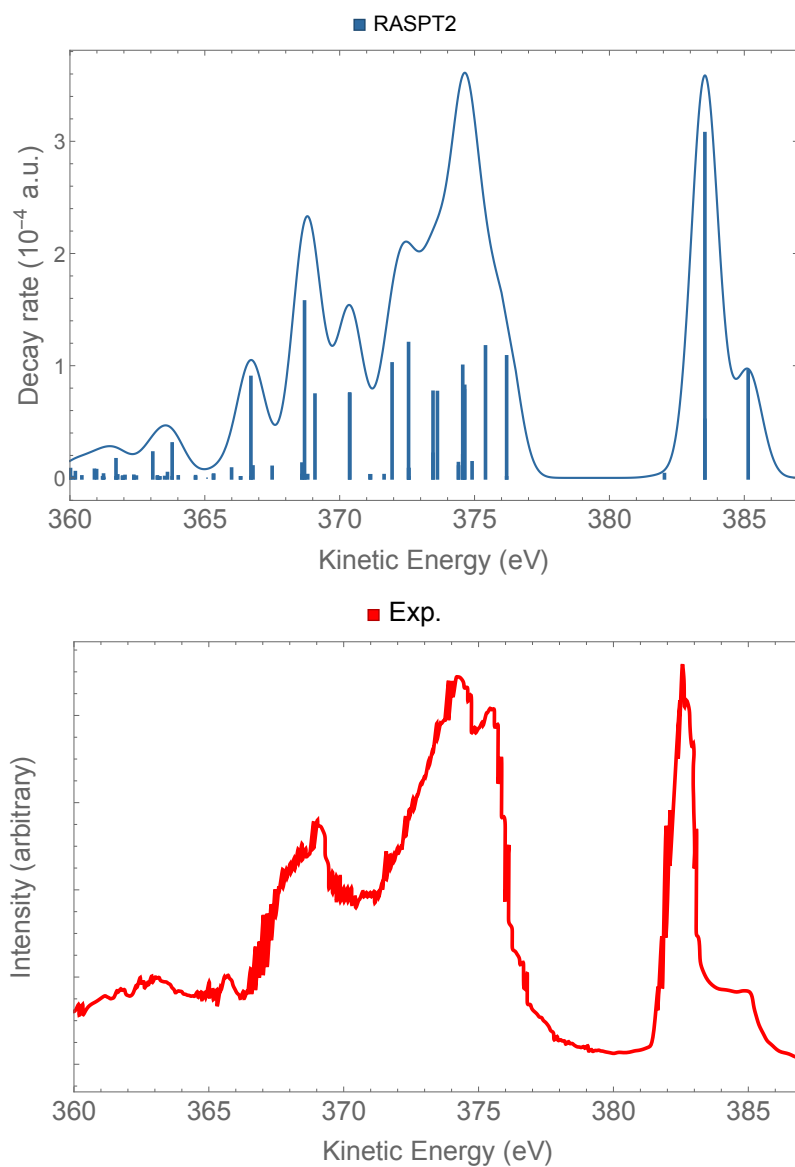


Figure S9: RAES spectrum of  $N_2$  convoluted with Gaussian functions using hwhm of 0.5 eV. The experimental points were extracted from ref. S30.

possible but at your own risk.

Additional input examples for normal Auger of water (singlet and triplet channels), RAES of N<sub>2</sub> without localization step, and RAES of CO at the carbon edge are found in the files `AES_H2O_singlet.input`, `AES_H2O_triplet.input`, `RAES_N2_c2v.input`, and `RAES_CO-C_edge.input` in the `SI_files` folder called `Auger`.

Examples of inputs for the CO and the H<sub>2</sub>O molecules, and their respective outputs, can also be found in the folder `Tools/AugerOca/auger_examples/`.

## S11 SA-SI-REKS Quantum–classical Trajectories

In this section additional information is provided with respect to the `OpenMolcas–GAMESS-US/Tinker` interface. In particular, the capabilities of the interface, the technical description of the interface, the description of rotary cycle in molecular motors, and results from MTDP simulations in solvent are described in detail.

### S11.1 Capabilities of the Interface

Here is reported the construction and application of a novel interface created to link `OpenMolcas` and the `GAMESS-US/Tinker` software packages. This interface enables SSR based simulations at the QM/MM level. More specifically, the interface exploits the availability of the SSR energy and analytical gradient in the latest version of `GAMESS-US`. Thus, by combining the molecular dynamics and surface hopping (`Dynamix` and `Surfacehop` modules, respectively) functionality of `OpenMolcas` with QM/MM energy and gradient calculations from `GAMESS-US/Tinker`, one enables the NAMD simulations using SSR level for the quantum mechanical treatment of photoresponsive molecular systems. The interface and the examples of calculations can be downloaded from GitLab repository.<sup>S31</sup>

The use of ensemble mixtures of electronic configurations leads to the occurrence of fractional occupations of frontier (molecular) orbitals, which are variationally optimized together with the molecular orbital coefficients. In the current version of `GAMESS-US`, the SSR method is implemented for systems with two fractionally occupied orbitals containing in total two active electrons. This is, in principle, sufficient to describe the ground and the first open-shell singlet excited states implicated in the  $\pi$ -bond torsion. The S<sub>0</sub> and S<sub>1</sub> states of such a system are represented in SSR formalism in terms of different electronic configurations. More specifically, the perfectly spin-paired singlet (PPS) and the open-shell singlet (OSS) configurations are included. In the SA-REKS(2,2) method, a weighted sum (an ensemble) of these configurations is self-consistently optimized. The energies of the individual states are then computed using the variationally optimized orbitals and fractional occupation numbers. The PPS energy is assumed to approximate the energy of the ground electronic state and the OSS energy approximates the excited electronic state energy. The SSR(2,2) method improves the description of the ground and excited states from SA-REKS(2,2) by including (possible) interaction between the PPS and OSS configuration. An extension of the SSR(2,2) method, the SSR(3,2) method is also available in `GAMESS-US`, where the configuration space is augmented by the doubly excited singlet (DES) configuration. For the SSR(2,2) and SSR(3,2) methods, `GAMESS-US` also provides the analytic energy gradient. In conclusion, the SSR(2,2) and the SSR(3,2) offer a performance that, potentially, parallel those of XMS-CASPT2/CASSCF(2,2) calculations of singlet states but provides the energies of only 2 electronic states and therefore



2 potential energy surfaces (PES). The necessary support of electrostatic potential fitted (ESPF) QM/MM calculations of the ESPF method has been already distributed within the `OpenMolcas/Tinker` interface and systematically used for the construction of QM/MM models.<sup>S32,S33</sup> Recently an improved version of ESPF<sup>S34</sup> was implemented in `GAMESS-US/Tinker` and is introduced in a local `GAMESS-US` build, while being anticipated to be included in the future official `GAMESS-US` releases. At the moment a patch enabling this non-standard version of `GAMESS-US` is available through the interface GitLab repository.

## S11.2 A Technical Description of the Interface

Below is a brief description of the interface. The `OpenMolcas` modules are used as the main driver for NAMD, while the `GAMESS-US/Tinker` is employed to calculate the SSR energies, CI coefficients and gradients. The handling of temporal files and parsing throughout the entire interface is implemented in shell scripts activated by the extended `MOLCAS` input language (EMIL) commands. The EMIL commands are specified in the `[Open]Molcas` input file. Each cycle of NAMD (equivalent to one time-step of MD) within the interface occurs in four stages (see the graphic representation in fig. S10):

*Stage 1.* Each NAMD cycle (i.e., the cycle connecting the current geometry and velocities to the next geometry and velocities) starts with `OpenMolcas` initializing shell scripts that update the electronic state driving the propagation, the geometry of the flexible part of the model at the current time-step ( $N$ ), and the molecular orbitals produced in the previous time-step ( $N - 1$ ) in `GAMESS/Tinker` input file that is used as a guess for the calculation of REKS energy and gradient at time-step  $N$ . The time difference between  $N - 1$  and  $N$  is typically 1 fs.

*Stage 2.* `GAMESS-US/Tinker` calculations with REKS-based QM/MM method are performed, which produces the potential energies of each electronic state in the geometry corresponding to time-step  $N$ , respective CI coefficients, energy gradient of the electronic state necessary for driving the propagation, and molecular orbitals. The shell scripts parse these data to the temporal files within the scratch folder.

*Stage 3.* The surface hopping module, `Surfacehop`, reads from the scratch folder the data necessary for Tully’s fewest switches surface hop algorithm, namely the electronic state driving the propagation, potential energies of each electronic state, and CI coefficients as produced by `GAMESS-US/Tinker`. After the execution of Tully’s fewest switches surface hopping algorithm, the `Surfacehop` module issues the decision on whether NAMD continues on the same PES or whether the system undergoes a hop between the two PESs, which leads to an update of the electronic state selected for NAMD propagation. If the hop is allowed, the EMIL command initiates another `GAMESS-US/Tinker` calculation to obtain the energy and gradient for the new NAMD state.

*Stage 4.* Molecular dynamics module, `Dynamix`, reads from the scratch folder the geometry of time-step  $N$ , the potential energies of each electronic state, and the energy gradient of the electronic state driving the propagation. Based on the REKS forces the Verlet algorithm generates the geometry of the time-step  $N + 1$ .

**Detailed Description of the Interface Workflow.** The initiation of the interface between `OpenMolcas` and `GAMESS-US/Tinker` requires: the `OpenMolcas` input file (`$Project.input`) specifying input parameters for `Dynamix` and `Surfacehop` modules of `OpenMolcas`; `RunFile` containing the information about the number of atoms and

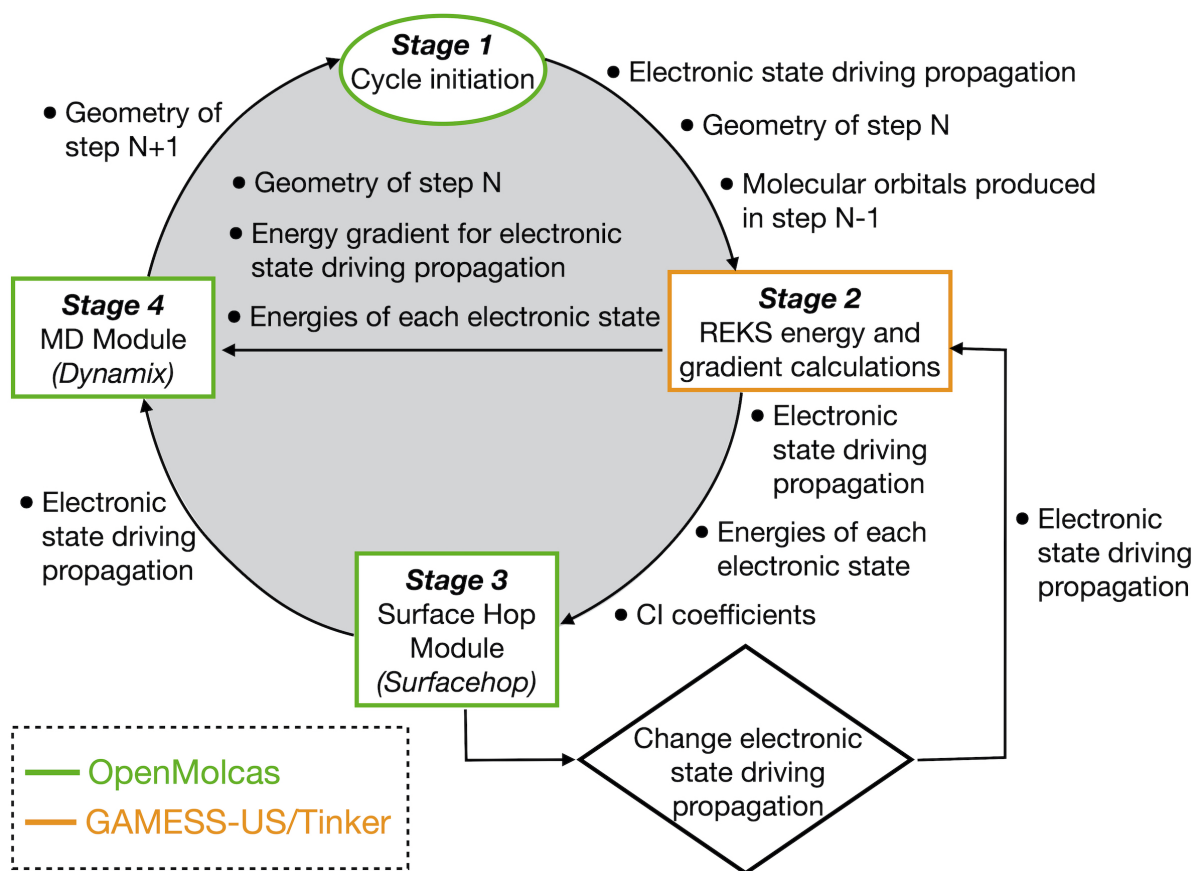


Figure S10: Schematic representation of the workflow in the OpenMolcas–GAMESS-US/Tinker interface.

atom types in the flexible part of the system; initial velocities file; JobIph file that allows for allocation of memory for CIBigArray of OpenMolcas; the input file specifying number of electronic states and CI coefficients (`$Project.in2`); as well as the GAMESS-US input file (`$Project.G.inp`) containing the geometry of the entire model and parameters relevant to QM calculations with SSR(2,2) and MM calculations with Tinker. The handling of temporal files and parsing throughout the entire interface is done with shell scripts that are activated with EMIL commands. The EMIL commands are specified in the OpenMolcas input file. All shell scripts are stored in the folder named GUTS. During the first cycle (first time step) in *Stage 1* the interface saves the number of electronic states, number of CI coefficients, the root number of current electronic state of NAMD propagation, and the geometry of the flexible part of the molecule in the scratch folder (OpenMolcas WorkDir). This task is performed by shell scripts `slice_in2.sh` and `slice_gamess.sh`, which extract this information from `$Project.in2` and `$Project.G.inp` files. In each subsequent cycle (each time step) shell scripts `update_gamess_geo.sh` and `update_gamess_vec.sh` insert current geometry of the flexible part of the molecule, the root number of current electronic state of NAMD propagation, and the molecular orbitals produced during previous cycle (which are used as the initial guess in current cycle) into `$Project.G.tmp.inp` file corresponding to the current time step.

In *Stage 2* the EMIL command `shell rungms-dev` initiates the GAMESS-US/Tinker calculations with SSR(2.2)-based QM/MM method, where GAMESS-US/Tinker produces potential energies corresponding to each electronic state, CI coefficients, energy gradient of current NAMD state, and molecular orbitals. To allow for parsing of produced data with OpenMolcas modules the EMIL commands execute the shell scripts that parse the corresponding data of GAMESS-US/Tinker output file to the temporal files within scratch folder, which is done in every cycle. These shell scripts are `get_gradient.sh`, `get_pot_energy.sh`, `get_S0_and_S1_energy.sh`, and `get_S0_and_S1_coeff.sh`. The temporal files corresponding to this data are `GRADIENT.txt`, `POT_ENERGY.txt`, `ROOTS_E.txt`, and `ROOTS_C.txt`.

In *Stage 3* the Surfacehop module of OpenMolcas reads from the scratch folder the data necessary for Tully fewest switches hop algorithm, namely the current state of NAMD, potential energies, and CI coefficients produced by GAMESS-US/Tinker. This is enabled through modifications of OpenMolcas source code, where Surfacehop module is instructed to read data from temporal files created with the shell scripts. After execution of Tully fewest switches surface hopping algorithm, the Surfacehop module issues the decision on whether the NAMD continues on the same PES or whether the system undergoes the surface hopping between two PES with subsequent update of NAMD state in the scratch folder. If surface hop is detected, the EMIL command initiates the GAMESS-US/Tinker calculations to obtain the energy gradient of new NAMD state. To achieve that we made a modification to a source code of Surfacehop module which upon surface hop creates temporal file `HOP.txt` and prints new root into the temporal file `RLXROOT.txt`. Using `IF (-FILE HOP.txt)` statement in OpenMolcas input and `slice_gamess_hop.sh` shell script, we update the NAMD state in `$Project.G.inp` file.

In final stage *Stage 4* the Dynamix module of OpenMolcas reads from the scratch folder the current geometry, the potential energies of each electronic state, and the energy gradient of current NAMD state, which were produced by GAMESS-US/Tinker. This is achieved with modifications to OpenMolcas code. Based on the forces calculated with SSR(2,2) the Verlet algorithm propagates the geometry of the system to the next time step and creates

temporal file `GEOMETRY.txt` in scratch folder, while shell script `update_gamess_geo.sh` writes this new geometry into `Project.G.tmp.inp`. Upon completion of *Stage 4* the cycle repeats for the next time step.

### S11.3 Description of Rotary Cycle in Molecular Motors

According to classic mechanism, absorption of the first photon results in breaking of the  $\pi$ -bond of the axle, which releases the strain energy and activates a counterclockwise (CCW) torsional motion of the rotor with respect to the stator. At nearly orthogonal configurations, the decay to the ground electronic state ( $S_0$ ) then occurs in the vicinity of a  $S_1/S_0$  conical intersection (CoIn) leading to restoration of the  $\pi$ -bond and relaxation to a stable intermediate with a Z configuration and an opposite M (minus) helicity (ZM). Then, the ZM intermediate is thermally converted to the final ZP product via a THI step making the system ready for a subsequent light-triggered Z to E half-rotation. After absorption of the second photon, a similar two-step CCW mechanism is activated, ultimately leading to the formation of the EM intermediate and reconstitution of the original EP reactant and completing the cycle.

### S11.4 Details of MTDP Simulations in Solvent

The QM/MM model and corresponding initial conditions were prepared using the methodology that was proven to be efficient in the simulations of molecular switches/motors and chromophores in protein and was discussed in detail in previous publications.<sup>S35-S40</sup> Initially the trajectories were executed for 850 fs at the SSR(2,2) level of theory. In four cases the propagation time was increased to 1100 fs since the hopping only occurred at times  $>790$  fs. After the hop from  $S_1$  to  $S_0$  each trajectory was propagated for an additional 200 fs using the same level of theory. At this point, the subset of “reactive” trajectories (i.e., the trajectories leading to the ZM photoproduct as opposed to the ones leading back to EP) were further propagated on the  $S_0$  PES using the faster KS DFT with BHHLYP functional allowing to describe the ZM to ZP transition occurring over picosecond time scale. The use of a less computationally expensive KS DFT method instead of SSR justified by the fact that, for the single-reference systems, both SSR and KS DFT yield essentially identical energies and gradients.<sup>S41,S42</sup> After the ZP was reached through a change in helicity, the trajectory was continued for 1 ps to ensure ZP stability. The total simulated times resulting from these calculations ranged from 1850 fs to 13 500 fs. Out of 40 simulated trajectories, 32 underwent surface hopping with the average  $S_1$  life-time of 482 fs. Of 32 trajectories that underwent surface hopping only 8 were reactive with surface hopping occurring at 396, 829, 799, 578, 343, 358, 385, and 370 fs, respectively. Six productive trajectories displayed a ZM to ZP transition that occurred at 1469, 4858, 648, 3249, 883 and 1324 fs, respectively. Two out of eight trajectories did not undergo ZM to ZP transition after the 13.5 ps of simulation.

## S12 Two-Step Acceleration of the Analytic Evaluation of the Vibronic Couplings for Multi-Root Systems

Input examples of the original and two-step procedure for the analytic evaluation of the vibronic couplings for multi-root systems are given in listings S40 and S41

Listing S40: OpenMolcas input tailored towards the original method.

```
1 &SEWARD
2 Title
3 Dy_bbpenCl_2020
4 ANGM
5 8.22323771761526982802 .23989506234286763540 8.02784003632896520089
6 AMFI
7
8 DOAN
9 RICD
10 SDIP
11
12 CDTH
13 1.0d-6
14
15 VERBOSE
16
17 Basis Set
18 DY.ANO-RCC-mb
19 DY1 4.351550 0.126947 4.248150 Angstrom
20 End of Basis Set
21
22 Basis Set
23 CL.ANO-RCC-mb
24 CL2 4.351550 -2.554927 4.248150 Angstrom
25 End of Basis Set
26
27 Basis Set
28 O.ANO-RCC-mb
29 O3 5.196621 0.609911 2.312693 Angstrom
30 End of Basis Set
31
32 Basis Set
33 N.ANO-RCC-mb
34 N4 5.763193 2.248941 4.695055 Angstrom
35 End of Basis Set
36
37 Basis Set
38 N.ANO-RCC-mb
39 N5 6.569970 -0.319559 5.486911 Angstrom
40 End of Basis Set
41
42 Basis Set
43 C.ANO-RCC-mb
44 C6 7.206167 -1.492994 5.580370 Angstrom
45 End of Basis Set
46
47 Basis Set
```

```
48 H.ANO-RCC-mb
49 H7 6.805824 -2.252377 5.174247 Angstrom
50 End of Basis Set
51
52 Basis Set
53 C.ANO-RCC-mb
54 C8 8.409806 -1.666518 6.234585 Angstrom
55 End of Basis Set
56
57 Basis Set
58 H.ANO-RCC-mb
59 H9 8.841479 -2.513522 6.246480 Angstrom
60 End of Basis Set
61
62 Basis Set
63 C.ANO-RCC-mb
64 C10 8.970285 -0.570396 6.870108 Angstrom
65 End of Basis Set
66
67 Basis Set
68 H.ANO-RCC-mb
69 H11 9.786636 -0.652863 7.349299 Angstrom
70 End of Basis Set
71
72 Basis Set
73 C.ANO-RCC-mb
74 C12 8.323645 0.625374 6.793641 Angstrom
75 End of Basis Set
76
77 Basis Set
78 H.ANO-RCC-mb
79 H13 8.702230 1.391629 7.211659 Angstrom
80 End of Basis Set
81
82 Basis Set
83 C.ANO-RCC-mb
84 C14 7.115655 0.742202 6.113937 Angstrom
85 End of Basis Set
86
87 Basis Set
88 C.ANO-RCC-mb
89 C15 6.391557 2.058236 6.052764 Angstrom
90 End of Basis Set
91
92 Basis Set
93 H.ANO-RCC-mb
94 H16 7.027753 2.795284 6.229487 Angstrom
95 End of Basis Set
96
97 Basis Set
98 H.ANO-RCC-mb
99 H17 5.689216 2.082289 6.749461 Angstrom
100 End of Basis Set
101
102 Basis Set
103 C.ANO-RCC-mb
104 C18 4.923344 3.480790 4.768124 Angstrom
105 End of Basis Set
```

106  
107 Basis Set  
108 H.ANO-RCC-mb  
109 H19 5.497748 4.274533 4.618589 Angstrom  
110 End of Basis Set  
111  
112 Basis Set  
113 H.ANO-RCC-mb  
114 H20 4.535185 3.554666 5.677228 Angstrom  
115 End of Basis Set  
116  
117 Basis Set  
118 C.ANO-RCC-mb  
119 C21 6.871968 2.401848 3.719680 Angstrom  
120 End of Basis Set  
121  
122 Basis Set  
123 H.ANO-RCC-mb  
124 H22 7.432447 1.585769 3.758763 Angstrom  
125 End of Basis Set  
126  
127 Basis Set  
128 H.ANO-RCC-mb  
129 H23 7.436799 3.161230 4.008554 Angstrom  
130 End of Basis Set  
131  
132 Basis Set  
133 C.ANO-RCC-mb  
134 C24 6.471625 2.628632 2.295700 Angstrom  
135 End of Basis Set  
136  
137 Basis Set  
138 C.ANO-RCC-mb  
139 C25 6.951166 3.716164 1.559921 Angstrom  
140 End of Basis Set  
141  
142 Basis Set  
143 H.ANO-RCC-mb  
144 H26 7.490758 4.369027 1.993232 Angstrom  
145 End of Basis Set  
146  
147 Basis Set  
148 C.ANO-RCC-mb  
149 C27 6.659612 3.862199 0.224302 Angstrom  
150 End of Basis Set  
151  
152 Basis Set  
153 H.ANO-RCC-mb  
154 H28 7.007736 4.600965 -0.259987 Angstrom  
155 End of Basis Set  
156  
157 Basis Set  
158 C.ANO-RCC-mb  
159 C29 5.865889 2.941319 -0.407822 Angstrom  
160 End of Basis Set  
161  
162 Basis Set  
163 H.ANO-RCC-mb

164 H30 5.649182 3.056429 -1.325423 Angstrom  
165 End of Basis Set  
166  
167 Basis Set  
168 C.ANO-RCC-mb  
169 C31 5.375905 1.843478 0.282077 Angstrom  
170 End of Basis Set  
171  
172 Basis Set  
173 H.ANO-RCC-mb  
174 H32 4.835442 1.209514 -0.175024 Angstrom  
175 End of Basis Set  
176  
177 Basis Set  
178 C.ANO-RCC-mb  
179 C33 5.662237 1.651056 1.627891 Angstrom  
180 End of Basis Set  
181  
182 Basis Set  
183 O.ANO-RCC-mb  
184 O34 3.506479 0.609911 6.183607 Angstrom  
185 End of Basis Set  
186  
187 Basis Set  
188 N.ANO-RCC-mb  
189 N35 2.939907 2.248941 3.801245 Angstrom  
190 End of Basis Set  
191  
192 Basis Set  
193 N.ANO-RCC-mb  
194 N36 2.133130 -0.319559 3.009389 Angstrom  
195 End of Basis Set  
196  
197 Basis Set  
198 C.ANO-RCC-mb  
199 C37 1.496933 -1.492994 2.915930 Angstrom  
200 End of Basis Set  
201  
202 Basis Set  
203 H.ANO-RCC-mb  
204 H38 1.897276 -2.252377 3.322053 Angstrom  
205 End of Basis Set  
206  
207 Basis Set  
208 C.ANO-RCC-mb  
209 C39 0.293294 -1.666518 2.261715 Angstrom  
210 End of Basis Set  
211  
212 Basis Set  
213 H.ANO-RCC-mb  
214 H40 -0.138379 -2.513522 2.249820 Angstrom  
215 End of Basis Set  
216  
217 Basis Set  
218 C.ANO-RCC-mb  
219 C41 -0.267185 -0.570396 1.626192 Angstrom  
220 End of Basis Set  
221



```
222 Basis Set
223 H.ANO-RCC-mb
224 H42 -1.083536 -0.652863 1.147001 Angstrom
225 End of Basis Set
226
227 Basis Set
228 C.ANO-RCC-mb
229 C43 0.379455 0.625374 1.702659 Angstrom
230 End of Basis Set
231
232 Basis Set
233 H.ANO-RCC-mb
234 H44 0.000870 1.391629 1.284641 Angstrom
235 End of Basis Set
236
237 Basis Set
238 C.ANO-RCC-mb
239 C45 1.587445 0.742202 2.382363 Angstrom
240 End of Basis Set
241
242 Basis Set
243 C.ANO-RCC-mb
244 C46 2.311543 2.058236 2.443536 Angstrom
245 End of Basis Set
246
247 Basis Set
248 H.ANO-RCC-mb
249 H47 1.675347 2.795284 2.266813 Angstrom
250 End of Basis Set
251
252 Basis Set
253 H.ANO-RCC-mb
254 H48 3.013884 2.082289 1.746839 Angstrom
255 End of Basis Set
256
257 Basis Set
258 C.ANO-RCC-mb
259 C49 3.779756 3.480790 3.728176 Angstrom
260 End of Basis Set
261
262 Basis Set
263 H.ANO-RCC-mb
264 H50 3.205352 4.274533 3.877711 Angstrom
265 End of Basis Set
266
267 Basis Set
268 H.ANO-RCC-mb
269 H51 4.167915 3.554666 2.819072 Angstrom
270 End of Basis Set
271
272 Basis Set
273 C.ANO-RCC-mb
274 C52 1.831132 2.401848 4.776620 Angstrom
275 End of Basis Set
276
277 Basis Set
278 H.ANO-RCC-mb
279 H53 1.270653 1.585769 4.737537 Angstrom
```

```
280 End of Basis Set
281
282 Basis Set
283 H.ANO-RCC-mb
284 H54 1.266301 3.161230 4.487746 Angstrom
285 End of Basis Set
286
287 Basis Set
288 C.ANO-RCC-mb
289 C55 2.231475 2.628632 6.200600 Angstrom
290 End of Basis Set
291
292 Basis Set
293 C.ANO-RCC-mb
294 C56 1.751934 3.716164 6.936379 Angstrom
295 End of Basis Set
296
297 Basis Set
298 H.ANO-RCC-mb
299 H57 1.212342 4.369027 6.503068 Angstrom
300 End of Basis Set
301
302 Basis Set
303 C.ANO-RCC-mb
304 C58 2.043488 3.862199 8.271998 Angstrom
305 End of Basis Set
306
307 Basis Set
308 H.ANO-RCC-mb
309 H59 1.695364 4.600965 8.756287 Angstrom
310 End of Basis Set
311
312 Basis Set
313 C.ANO-RCC-mb
314 C60 2.837211 2.941319 8.904122 Angstrom
315 End of Basis Set
316
317 Basis Set
318 H.ANO-RCC-mb
319 H61 3.053918 3.056429 9.821723 Angstrom
320 End of Basis Set
321
322 Basis Set
323 C.ANO-RCC-mb
324 C62 3.327195 1.843478 8.214223 Angstrom
325 End of Basis Set
326
327 Basis Set
328 H.ANO-RCC-mb
329 H63 3.867658 1.209514 8.671324 Angstrom
330 End of Basis Set
331
332 Basis Set
333 C.ANO-RCC-mb
334 C64 3.040863 1.651056 6.868409 Angstrom
335 End of Basis Set
336
337 End of Input
```

```
338
339
340
341
342 >>COPY $FileDir/Dy_bbpenCl_2020_mb_mb_cas_9in7_ms6.RasOrb $WorkDir/INPORB
343 &RASSCF
344 LUMO
345 SPIN
346 6
347 NACTEL
348 9 0 0
349 FROZ
350 0
351 INAC
352 157
353 RAS1
354 0
355 RAS2
356 7
357 RAS3
358 0
359 CIROOT
360 21 21 1
361 ITERations
362 200 20
363 CIMX
364 20
365 ORBL
366 ALL
367 ORBA
368 FULL
369 End Of Input
370
371
372 &MCLR
373 ITER
374 200
375 EXPD
376 21
377 PRINT
378 3
379 THREshold
380 1.0d-10
381 NAC
382 1 5
383 End Of Input
384
385
386 &ALASKA
387 SHOW
388 VERBOSE
389 CUTOFF
390 1.0d-14
391 PNEW
392 NAC
393 1 5
394 End Of Input
```

Listing S41: OpenMolcas input tailored towards the new two-step approach.

```

1 &SEWARD
2 Title
3 Dy_bbpenCl_2020
4 ANGM
5 8.22323771761526982802 .23989506234286763540 8.02784003632896520089
6 AMFI
7
8 DOAN
9 RICD
10 SDIP
11
12 CDTH
13 1.0d-6
14
15 VERBOSE
16
17 Basis Set
18 DY.ANO-RCC-mb
19 DY1 4.351550 0.126947 4.248150 Angstrom
20 End of Basis Set
21
22 Basis Set
23 CL.ANO-RCC-mb
24 CL2 4.351550 -2.554927 4.248150 Angstrom
25 End of Basis Set
26
27 Basis Set
28 O.ANO-RCC-mb
29 O3 5.196621 0.609911 2.312693 Angstrom
30 End of Basis Set
31
32 Basis Set
33 N.ANO-RCC-mb
34 N4 5.763193 2.248941 4.695055 Angstrom
35 End of Basis Set
36
37 Basis Set
38 N.ANO-RCC-mb
39 N5 6.569970 -0.319559 5.486911 Angstrom
40 End of Basis Set
41
42 Basis Set
43 C.ANO-RCC-mb
44 C6 7.206167 -1.492994 5.580370 Angstrom
45 End of Basis Set
46
47 Basis Set
48 H.ANO-RCC-mb
49 H7 6.805824 -2.252377 5.174247 Angstrom
50 End of Basis Set
51
52 Basis Set
53 C.ANO-RCC-mb
54 C8 8.409806 -1.666518 6.234585 Angstrom
55 End of Basis Set
56
57 Basis Set

```

```
58 H.ANO-RCC-mb
59 H9 8.841479 -2.513522 6.246480 Angstrom
60 End of Basis Set
61
62 Basis Set
63 C.ANO-RCC-mb
64 C10 8.970285 -0.570396 6.870108 Angstrom
65 End of Basis Set
66
67 Basis Set
68 H.ANO-RCC-mb
69 H11 9.786636 -0.652863 7.349299 Angstrom
70 End of Basis Set
71
72 Basis Set
73 C.ANO-RCC-mb
74 C12 8.323645 0.625374 6.793641 Angstrom
75 End of Basis Set
76
77 Basis Set
78 H.ANO-RCC-mb
79 H13 8.702230 1.391629 7.211659 Angstrom
80 End of Basis Set
81
82 Basis Set
83 C.ANO-RCC-mb
84 C14 7.115655 0.742202 6.113937 Angstrom
85 End of Basis Set
86
87 Basis Set
88 C.ANO-RCC-mb
89 C15 6.391557 2.058236 6.052764 Angstrom
90 End of Basis Set
91
92 Basis Set
93 H.ANO-RCC-mb
94 H16 7.027753 2.795284 6.229487 Angstrom
95 End of Basis Set
96
97 Basis Set
98 H.ANO-RCC-mb
99 H17 5.689216 2.082289 6.749461 Angstrom
100 End of Basis Set
101
102 Basis Set
103 C.ANO-RCC-mb
104 C18 4.923344 3.480790 4.768124 Angstrom
105 End of Basis Set
106
107 Basis Set
108 H.ANO-RCC-mb
109 H19 5.497748 4.274533 4.618589 Angstrom
110 End of Basis Set
111
112 Basis Set
113 H.ANO-RCC-mb
114 H20 4.535185 3.554666 5.677228 Angstrom
115 End of Basis Set
```

```
116
117 Basis Set
118 C.ANO-RCC-mb
119 C21 6.871968 2.401848 3.719680 Angstrom
120 End of Basis Set
121
122 Basis Set
123 H.ANO-RCC-mb
124 H22 7.432447 1.585769 3.758763 Angstrom
125 End of Basis Set
126
127 Basis Set
128 H.ANO-RCC-mb
129 H23 7.436799 3.161230 4.008554 Angstrom
130 End of Basis Set
131
132 Basis Set
133 C.ANO-RCC-mb
134 C24 6.471625 2.628632 2.295700 Angstrom
135 End of Basis Set
136
137 Basis Set
138 C.ANO-RCC-mb
139 C25 6.951166 3.716164 1.559921 Angstrom
140 End of Basis Set
141
142 Basis Set
143 H.ANO-RCC-mb
144 H26 7.490758 4.369027 1.993232 Angstrom
145 End of Basis Set
146
147 Basis Set
148 C.ANO-RCC-mb
149 C27 6.659612 3.862199 0.224302 Angstrom
150 End of Basis Set
151
152 Basis Set
153 H.ANO-RCC-mb
154 H28 7.007736 4.600965 -0.259987 Angstrom
155 End of Basis Set
156
157 Basis Set
158 C.ANO-RCC-mb
159 C29 5.865889 2.941319 -0.407822 Angstrom
160 End of Basis Set
161
162 Basis Set
163 H.ANO-RCC-mb
164 H30 5.649182 3.056429 -1.325423 Angstrom
165 End of Basis Set
166
167 Basis Set
168 C.ANO-RCC-mb
169 C31 5.375905 1.843478 0.282077 Angstrom
170 End of Basis Set
171
172 Basis Set
173 H.ANO-RCC-mb
```

174 H32 4.835442 1.209514 -0.175024 Angstrom  
175 End of Basis Set  
176  
177 Basis Set  
178 C.ANO-RCC-mb  
179 C33 5.662237 1.651056 1.627891 Angstrom  
180 End of Basis Set  
181  
182 Basis Set  
183 O.ANO-RCC-mb  
184 O34 3.506479 0.609911 6.183607 Angstrom  
185 End of Basis Set  
186  
187 Basis Set  
188 N.ANO-RCC-mb  
189 N35 2.939907 2.248941 3.801245 Angstrom  
190 End of Basis Set  
191  
192 Basis Set  
193 N.ANO-RCC-mb  
194 N36 2.133130 -0.319559 3.009389 Angstrom  
195 End of Basis Set  
196  
197 Basis Set  
198 C.ANO-RCC-mb  
199 C37 1.496933 -1.492994 2.915930 Angstrom  
200 End of Basis Set  
201  
202 Basis Set  
203 H.ANO-RCC-mb  
204 H38 1.897276 -2.252377 3.322053 Angstrom  
205 End of Basis Set  
206  
207 Basis Set  
208 C.ANO-RCC-mb  
209 C39 0.293294 -1.666518 2.261715 Angstrom  
210 End of Basis Set  
211  
212 Basis Set  
213 H.ANO-RCC-mb  
214 H40 -0.138379 -2.513522 2.249820 Angstrom  
215 End of Basis Set  
216  
217 Basis Set  
218 C.ANO-RCC-mb  
219 C41 -0.267185 -0.570396 1.626192 Angstrom  
220 End of Basis Set  
221  
222 Basis Set  
223 H.ANO-RCC-mb  
224 H42 -1.083536 -0.652863 1.147001 Angstrom  
225 End of Basis Set  
226  
227 Basis Set  
228 C.ANO-RCC-mb  
229 C43 0.379455 0.625374 1.702659 Angstrom  
230 End of Basis Set  
231

```
232 Basis Set
233 H.ANO-RCC-mb
234 H44 0.000870 1.391629 1.284641 Angstrom
235 End of Basis Set
236
237 Basis Set
238 C.ANO-RCC-mb
239 C45 1.587445 0.742202 2.382363 Angstrom
240 End of Basis Set
241
242 Basis Set
243 C.ANO-RCC-mb
244 C46 2.311543 2.058236 2.443536 Angstrom
245 End of Basis Set
246
247 Basis Set
248 H.ANO-RCC-mb
249 H47 1.675347 2.795284 2.266813 Angstrom
250 End of Basis Set
251
252 Basis Set
253 H.ANO-RCC-mb
254 H48 3.013884 2.082289 1.746839 Angstrom
255 End of Basis Set
256
257 Basis Set
258 C.ANO-RCC-mb
259 C49 3.779756 3.480790 3.728176 Angstrom
260 End of Basis Set
261
262 Basis Set
263 H.ANO-RCC-mb
264 H50 3.205352 4.274533 3.877711 Angstrom
265 End of Basis Set
266
267 Basis Set
268 H.ANO-RCC-mb
269 H51 4.167915 3.554666 2.819072 Angstrom
270 End of Basis Set
271
272 Basis Set
273 C.ANO-RCC-mb
274 C52 1.831132 2.401848 4.776620 Angstrom
275 End of Basis Set
276
277 Basis Set
278 H.ANO-RCC-mb
279 H53 1.270653 1.585769 4.737537 Angstrom
280 End of Basis Set
281
282 Basis Set
283 H.ANO-RCC-mb
284 H54 1.266301 3.161230 4.487746 Angstrom
285 End of Basis Set
286
287 Basis Set
288 C.ANO-RCC-mb
289 C55 2.231475 2.628632 6.200600 Angstrom
```



```

290 End of Basis Set
291
292 Basis Set
293 C.ANO-RCC-mb
294 C56 1.751934 3.716164 6.936379 Angstrom
295 End of Basis Set
296
297 Basis Set
298 H.ANO-RCC-mb
299 H57 1.212342 4.369027 6.503068 Angstrom
300 End of Basis Set
301
302 Basis Set
303 C.ANO-RCC-mb
304 C58 2.043488 3.862199 8.271998 Angstrom
305 End of Basis Set
306
307 Basis Set
308 H.ANO-RCC-mb
309 H59 1.695364 4.600965 8.756287 Angstrom
310 End of Basis Set
311
312 Basis Set
313 C.ANO-RCC-mb
314 C60 2.837211 2.941319 8.904122 Angstrom
315 End of Basis Set
316
317 Basis Set
318 H.ANO-RCC-mb
319 H61 3.053918 3.056429 9.821723 Angstrom
320 End of Basis Set
321
322 Basis Set
323 C.ANO-RCC-mb
324 C62 3.327195 1.843478 8.214223 Angstrom
325 End of Basis Set
326
327 Basis Set
328 H.ANO-RCC-mb
329 H63 3.867658 1.209514 8.671324 Angstrom
330 End of Basis Set
331
332 Basis Set
333 C.ANO-RCC-mb
334 C64 3.040863 1.651056 6.868409 Angstrom
335 End of Basis Set
336
337 End of Input
338
339
340
341
342 >>COPY $FileDir/Dy_bbpenCl_2020_mb_mb_cas_9in7_ms6.RasOrb $WorkDir/INPORB
343 &RASSCF
344 LUMORB
345 SPIN
346 6
347 NACTEL

```

```
348 9 0 0
349 FROZ
350 0
351 INAC
352 157
353 RAS1
354 0
355 RAS2
356 7
357 RAS3
358 0
359 CIROOT
360 21 21 1
361 ITERations
362 200 20
363 CIMX
364 20
365 ORBL
366 ALL
367 ORBA
368 FULL
369 End Of Input
370
371
372
373 &MCLR
374 ITER
375 200
376 EXPD
377 21
378 PRINT
379 3
380 THREshold
381 1.0d-10
382 SALA
383 1
384 TWOSTep
385 FIRST
386 End Of Input
387
388
389 &MCLR
390 ITER
391 200
392 EXPD
393 21
394 PRINT
395 3
396 THREshold
397 1.0d-10
398 NAC
399 1 5
400 TWOSTep
401 SECOND
402 End Of Input
403
404
405 &ALASKA
```

```
406 SHOW
407 VERBOSE
408 CUTOFF
409 1.0d-14
410 PNEW
411 NAC
412 1 5
413 End Of Input
```

## S13 MC-PDFT Nonadiabatic Dynamics

The SHARC–OpenMolcas interface for MC-PDFT is available in the MCPDFT directory with example input and output files. In addition to replacing the interface script, the user must define the method as *MCPDFT* and provide the functional (e.g., *t:PBE*) and grid size (*coarse*, *sg1fine*, *fine*, or *ultrafine*) in the `MOLCAS.template` file. This interface script can also run simulations with CMS-PDFT by defining the method as *CMS-PDFT*, although this has not been well tested yet.

## S14 ANO-R Basis Set

Tools for basis set generation.

### S14.1 Hartree–Fock energies for atoms with higher ANO-R-*N* basis sets

The convergence of the results with an increase of the basis set is essential test for consistency of the basis set. Many different extrapolation formulae exist,<sup>S43–S45</sup> all based on various empirical evaluations of basis set convergence. Regarding ANO-type of the basis sets, as it was pointed in the literature,<sup>S46,S47</sup> the standard contraction levels (e.g., ANO-L-VDZP, ANO-L-VTZP, ANO-L-VQZP) are not necessarily follow the extrapolation pattern.

Instead, one may exploit the fact that for ANO-type basis sets, any combination of contracted functions for a given element can be used.<sup>S48,S49</sup> To clarify, for C, there are 9 s, 7 p, 3 d and 2 f contracted basis functions in the ANO-R basis set. Thus, the largest ANO-R basis set for C is 9s7p4d2f, much larger than ANO-R-2 (4s4p2d1f), which was chosen as the largest standard contraction level for C based on the observations in the original ANO-R paper.<sup>S50</sup> By removing one shell of functions at a time from the full basis set, basis sets which allow for a smooth basis set extrapolation is obtained. Here, we generated four such basis sets for each element. For C, this means 9s7p3d2f, 8s6p2d1f, 7s5p1d and 6s4p.

Figure S11 shows the convergence of the Hartree–Fock energies for atoms with an increase of the basis set from ANO-R-1 to ANO-R-8. The higher than R-3 contraction schemes were made by increasing by “layers” of s-, p-, d- etc. orbitals.

Atomic states used in the calculations in this project taken from ref. S51. The biggest challenge for such calculations is to ensure that the converged results corresponds to the same electron configuration. The latter can be achieved by restarting the calculations from the the previous calculation with smaller basis set. For contractions higher than 3

the changes in the total energies are very small, and for the majority of the elements they are less than  $10^{-6} E_h$ .

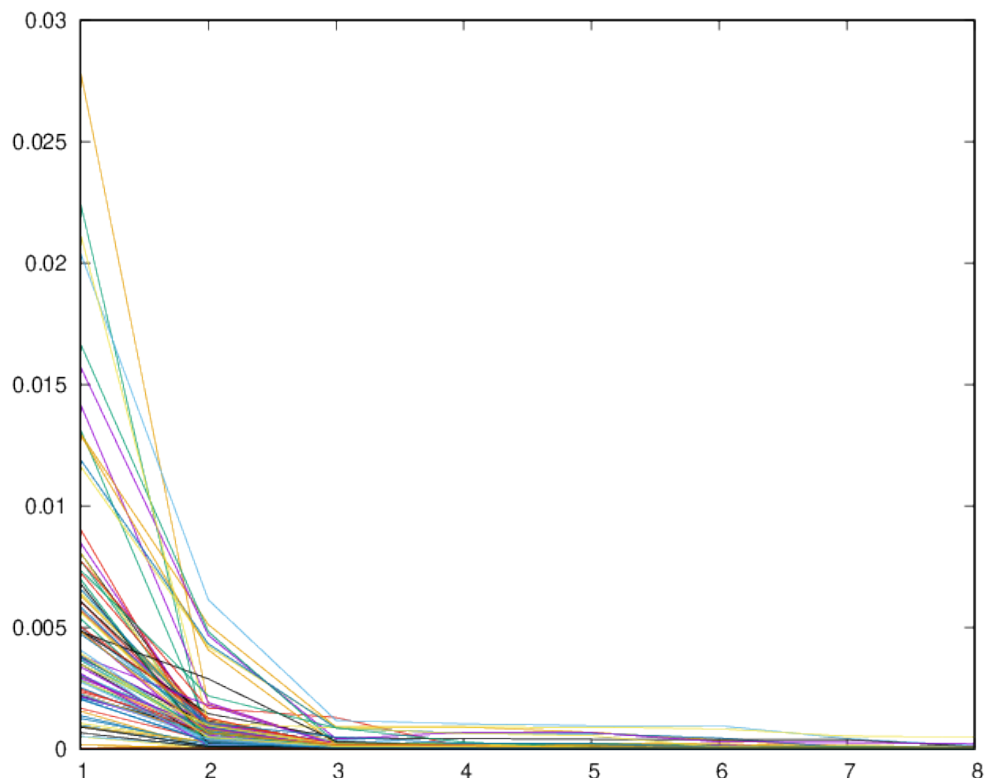


Figure S11: Convergence of Hartree–Fock atomic energies (from H to Rn) with an increase of the basis set quality. The  $x$  axis corresponds to ANO-R- $N$  ( $N=1,2,3,\dots,8$ ), the  $y$  axis shows the error in  $E_h$ .

## S14.2 Electron affinity of oxygen atom

Electron affinity for oxygen atom are computed as a difference between CASPT2 energies for the ground (triplet) state of oxygen and anion  $O^-$ . Active space used in the calculations include  $1s2s2p3s3p$  orbitals. The further increase of the active space (to 14 orbitals) leads to a minor change of the results with the order of 0.02 eV.

## S15 Use of SCEPIC program for generation of AIMPs and for creating input templates for OpenMolcas calculations

This section describes two possible use of SCEPIC program: Generation of ab initio model potentials (AIMP) for a new crystal and generation of input for OpenMolcas with the use of AIMPs.

Abbreviations used in the section include the quantum part (QP), potential part (PP), and point charges (CP). The QP is set by the coordinate file in XYZ format.

Table S17: Electron affinity of oxygen, in eV, computed with different ANO-R basis sets

$N$	=	3	4	5	6	7	8	9	10
$Ns(N-2)p1d$		-7.89	-0.60	0.31	0.67	0.93	0.99	0.99	0.99
$Ns(N-1)p1d$		-0.77	0.27	0.66	0.93	0.99	0.99	0.99	
$NsNp1d$		0.13	0.62	0.92	0.99	0.99	0.99		
$Ns(N-2)p2d$		-7.82	-0.47	0.44	0.80	1.04	1.10	1.10	1.10
$NsNp2d$		0.24	0.75	1.04	1.10	1.10	1.10		
$Ns(N-2)p3d1f$		-7.68	-0.34	0.55	0.90	1.14	1.19	1.19	1.19
$Ns(N-2)p4d2f$		-7.65	-0.29	0.59	0.94	1.18	1.23	1.23	1.23
$Ns(N-2)p4d2f1g$		-7.59	-0.22	0.64	0.98	1.22	1.27	1.27	1.27
$Ns(N-2)p4d3f2g1h$		-7.56	-0.19	0.67	1.02	1.25	1.30	1.30	1.30

## S16 Ab Initio Model Potentials

Generation of AIMP potentials for a new crystal.

### S16.1 Quick start

Ab initio model potentials should be generated for any new crystal structure. This section gives a step-by-step tutorial for this purpose. One option is to create input files for SCEPIC code directly, but for simple cases, it is possible to generate these inputs by `sceplic_inp.plx` code, which can be downloaded from SCEPIC homepage (<https://www.molcas.org/SCEPIC>).

One needs to create an XYZ file for your periodic structure, with an extended MOLCAS format. An example of such a format is generated by running `sceplic_inp.plx` code with a flag `-h`. The file should contain all atoms in the unit cell but in addition to that,

- it should contain a translation section at the end in the form of four lines, e.g.:

```
1 #Origin 0 0 0
2 #Vector1 2.0 0 0
3 #Vector2 0 2.0 0
4 #Vector3 0 0 2.0
```

- Atoms for which AIMP potentials should be constructed had additional parameters: charge and orbitals. The orbitals are given by letters (no spaces)

```
1 Mg 0 0 0 +2 sp
```

Note that Cartesian coordinates in the file should be inside the unit cell, which is defined by the vectors, so the partial coordinates are in the range  $[0, 1]$ .

In case one atom is located at different Wyckoff positions, the element name must be labeled, by adding an underscore sign and any label, e.g.,

```
1 Mg_a 0 0 0 +2 sp
```

After creating the coordinate file, execute `./sceplic_inp.plx filename.xyz`. The script will create input files for SCEPIC (small scripts, written in Julia language). These scripts can be modified for more efficient use of SCEPIC program.

## S16.2 Drivers for embedding script

There are two alternative ways to set up the size of the CP: by radius (*Charges\_Radius*) or by extension of the unit cells (*Charges\_Size*). These keywords are exclusive. The *Charges\_Radius* keyword requires one number. It can be complemented by keyword *Charges\_Center*, which defines the center of the sphere. If *Charges\_Size* is defined instead, it requires one, three, or six numbers, defining the translation of the unit cell (uniform in all directions, symmetrical, or arbitrary). The setup of the AIMP region is similar to the setup of CP. Two alternative keywords can be used: *AIMP\_Radius* or *AIMP\_Size*.

The complete example of `driver.jl` script:

```
1 using embedding_factory
2
3 # define the coordinates of the system
4 #
5 # Unit cell file (it must contain in the comment line LATTICE followed by 9 numbers
6 uc_file = "./MgO_bulk.xyz"
7 # Quantum part (ordinary xyz file)
8 qm_file = "./MgO_clust.xyz"
9 #definition of elements and charges
10 charge_dict = Dict("Mg" => 2, "O" => -2)
11
12 # region of charges
13 Charges_Radius=20
14 Charges_Center=[0.0,0.0,0.0]
15 #Charges_Size=[4, 4, 4]
16
17 # region of AIMPs
18 AIMP_Radius=5
19 #AIMP_Size=[2, 2, 2]
20
21
22 #Multipoles to eliminate. Max=10, Default=4
23 MaxMultipole=3
24
25 #template for Spherical based cluster
26 input = Dict("elmoment" => MaxMultipole,
27             "unit_cell_file" => uc_file,
28             "qm_file" => qm_file,
29             "cluster_radius" => Charges_Radius,
30             "aimp_radius" => AIMP_Radius,
31             "reference" => Charges_Center,
32             "charge_dict" => charge_dict)
33
34 #template for Unit cell based cluster
35 input_UC = Dict("elmoment" => MaxMultipole,
36               "unit_cell_file" => uc_file,
37               "qm_file" => qm_file,
38               "cluster_size" => Charges_Size,
39               "aimp_radius" => AIMP_Size,
40               "reference" => Charges_Center,
41               "charge_dict" => charge_dict)
42 embedding(input)
43 #embedding(input_UC)
```

### S16.3 Driver for embedding

For each atom in the system, one needs to prepare `xfield` file and `aimpfield` file. In order to do that we have to run individual calculations for each atom. The lattice coordinate file is the same as in the previous example, but the quantum part contains only one single atom. In the case of Mg, since it is located at (0,0,0), the only modification we need is to change the coordinate file, keeping only one atom in it. To avoid confusion, you can copy the coordinate file, edit it and change the name of the file in the driver script.

In the driver file, we have to set *Charges\_Radius* and *AIMP\_Radius* to be large enough. *AIMP\_Radius* must be at least 6 Å, and for the Charges – the electrostatic potential should converge.

Run `julia driver.jl` command. Rename `project.xfield` to `Mg.xfield`. Rename `project.inp` to `Mg.aimpfield` and remove from the top of the file the first 4 lines (the basis set for the real atom).

Next, we need to repeat the procedure for O atom. The coordinate file is:

```
1 1
2 comment
3 0 2.128242 0 0
```

Since O is not in the center of the coordinates, we have to change the center of the sphere in the driver script:

```
1 qm_file = "./O_clust.xyz"
2 Charges_Radius=40
3 Charges_Center=[2.1282420,0.0,0.0]
4 AIMP_Radius=10
```

Run `julia driver-emb.jl` and as before rename `project.xfield` to `O.xfield` and `project.inp` to `O.aimpfield` (remove 4 first lines).

### S16.4 Driver script for SCEPIC

At the beginning of the script, there is a definition for each ion, which include the following parameters: element name, position (should match the corresponding selection in embedding input, formal atomic charge, exponents for valence shells (s- exponents are compulsory, others are optional), relativistic hamiltonian, nucleus, indication of Restricted or Unrestricted Hartree–Fock hamiltonian, DFT functional, and links to the files, generated by embedding program.

The exponents can be found in the basis library file. For UHF hamiltonian, the Multiplicity must be set.

Example of the driver file for MgO

```
1 using sceplic_module
2
3 Mg_s = [169973.87,24699.066,7175.4851,2478.7920,917.73259,351.75140,
4         137.51557,54.430286,21.722569,8.7194484,3.5147166,1.4212132,
5         .57607650,.23395250,.09515600,.03875100,.01550040]
6 Mg_p = [557.13078,135.48388,47.212958,18.000955,7.1146607,2.8630240,
7         1.1640185,.47630470,.19573560,.08067830,.03332660,.01333060]
8 ion1 = Dict("element" => "Mg",
9            "position" => [ 0.00000000, 0.00000000, 0.00000000],
10           "charge" => +2,
11           "s_basis" =>Mg_s,
```

```

12     "p_basis" => Mg_p,
13     "hamiltonian" => "rx2c",
14     "nucleus" => "finite",
15     "SCF" => "RHF",
16     "ksdft" => "PBE",
17     "AIMPLabel" => "Mg.ECP.author.0s.0s.0e-update",
18     "AIMPEmbedding" => "./Mg.aimpfield",
19     "xfield" => "./Mg.xfield")
20
21 O_s = [105374.95, 15679.240, 3534.5447, 987.36516, 315.97875, 111.65428,
22       42.699451, 17.395596, 7.4383090, 3.2228620, 1.2538770, .49515500,
23       .19166500, .06708300]
24 O_p = [200.00000, 46.533367, 14.621809, 5.3130640, 2.1025250, .85022300,
25       .33759700, .12889200, .04511200]
26 ion2 = Dict("element" => "O",
27            "position" => [ 2.12824200, 0.00000000, 0.00000000],
28            "charge" => -2,
29            "s_basis" => O_s,
30            "p_basis" => O_p,
31            "hamiltonian" => "rx2c",
32            "nucleus" => "finite",
33            "SCF" => "RHF",
34            "ksdft" => "PBE",
35            "AIMPLabel" => "O.ECP.author.0s.0s.0e-update",
36            "AIMPEmbedding" => "./O.aimpfield",
37            "xfield" => "./O.xfield")
38
39 ion_dict = [ion1, ion2]
40 input = Dict("maxiter" => 20)
41
42 sceplic(ion_dict)#, input)

```

Executing this script will generate ECP file to be appended to AIMP file in `basis_library` directory of OpenMolcas.

## S16.5 Generation of input for OpenMolcas calculation with AIMPs

SCEPIC (or, more precisely, the embedding part of SCEPIC code) can be used for the generation of a template for OpenMolcas input. One should start from a selection of the quantum part of the cluster. The cluster can have a charge, which later will be compensated by electrostatic field. The positively charged clusters are preferred.

At this step it is important that all atoms in the QC part are located in the exact positions as in the ideal crystal. After the generation of the inputs, the atoms in the QC part can be moved, e.g., accordingly to optimized geometry.

## S16.6 Diffusion of oxygen ad-atom on MgO(001)

In this section, the usage of AIMPs for surface reactions will be discussed, exemplified by the diffusion of an oxygen ad-atom on the MgO(001) surface. This reaction has previously been proposed to be an example of a spin-crossing, where the lowest energy route would correspond to changing between singlet and triplet spin-states.<sup>S52</sup> Spin-crossing reactions are typical examples of where multiconfigurational approaches are important, for which reason we reexamined this reaction using an AIMP embedding.



To generate the geometries for the reaction, the following steps were followed. First, the geometries were optimized by periodic DFT. Thereafter, an embedding environment was generated for a pristine MgO(001) surface; in the production calculations, the periodic DFT geometries were pasted into this embedding. Since AIMP are, in principle at least, sensitive to the position in which it was optimized, two sets of AIMP were used, one to represent bulk ions and one to represent surface ions. This can be achieved in SCEPIC by first generating surface embedding environments, which require that *surface => true* in the *embedding* code to generate an appropriate embedding environment (the code assumes that the surface normal is parallel to the third unit-cell vector) and then optimizing a set of bulk AIMP, which are then frozen in the optimization of surface AIMP, using the keyword *external\_AIMP*. Note that in this case, no coordinates in the QM region can be larger than 1.0 in the internal coordinates of the unit-cell. Once these steps have been followed, typical AIMP calculations can be made in *OpenMolcas*.

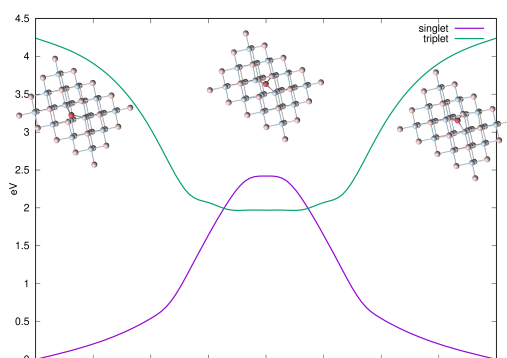


Figure S12: Diffusion add-atom oxygen on the surface of MgO

A preliminary set of results for this diffusion reaction is presented in fig. S12. These were obtained by performing DMRGSCF (*qCMAQUIS*) calculations, correlating the 2s2p3s3p orbitals of the three oxygens involved in the diffusion step. Preliminary calculations were made with a fairly small basis set, x2c-SVPALL with the timing for integrals being around 24 h for an Mg<sub>8</sub>O<sub>20</sub>-cluster (plus the adsorbed oxygen) and the DMRGSCF iterations similarly being around 24 h. An example input is presented below. On the border AIMP, orthogonality functions have been placed.<sup>S53</sup>

```

1  &GATEWAY
2  Basis set
3  Mg. X2C-SVPA11
4  Mg1      0.0431225000    2.2225390000    2.2235268571    Angstrom
5  Mg2      2.2226725000    0.0432160000    2.2237518571    Angstrom
6  Mg3     -2.2551525000    -0.0009020000    2.1418498571    Angstrom
7  Mg4     -0.0009145000    -2.2551750000    2.1418698571    Angstrom
8  Mg5      2.1787395000    4.3178470000    2.1031268571    Angstrom
9  Mg6      4.3176005000    2.1783730000    2.0683868571    Angstrom
10 Mg7      0.0012615000    0.0012900000    -0.0335731429    Angstrom
11 Mg8      2.1418855000    2.1425770000    0.0090408571    Angstrom
12 End of Basis
13 Basis set
14 O. X2C-SVPA11
15 O1      0.4747975000    0.4748860000    3.2619938571    Angstrom
16 O2      2.1325085000    -2.1395600000    2.1531948571    Angstrom
17 O3     -2.1397605000    2.1323460000    2.1530178571    Angstrom
18 O4      0.0034335000    4.3110760000    2.1450798571    Angstrom

```

```

19      05          4.3112325000    0.0035180000    2.1380178571  Angstrom
20      06         -2.1804605000   -2.1804810000    2.1358688571  Angstrom
21      07          4.3126935000    4.3121300000    2.1151308571  Angstrom
22      08         -4.3070655000    0.0002130000    2.1187428571  Angstrom
23      09          0.0002465000   -4.3070260000    2.1187428571  Angstrom
24      010         2.1476645000    6.4336420000    2.1186868571  Angstrom
25      011         6.4331045000    2.1473620000    2.1111888571  Angstrom
26      012        -0.2125775000   -0.2125660000    2.0809038571  Angstrom
27      013         2.2283395000    2.2274170000    2.0849338571  Angstrom
28      014         0.0095665000    2.1229320000    0.0476488571  Angstrom
29      015         2.1229595000    0.0102300000    0.0473448571  Angstrom
30      016        -0.0029405000   -2.1389070000    0.0064948571  Angstrom
31      017        -2.1389785000   -0.0030590000    0.0064398571  Angstrom
32      018         2.1428335000    4.2991510000   -0.0140371429  Angstrom
33      019         4.2991175000    2.1427510000   -0.0204941429  Angstrom
34      020         0.0026455000    0.0026820000   -2.1020091429  Angstrom
35      021         2.1357225000    2.1358840000   -2.0948771429  Angstrom
36      End of Basis
37      Basis set
38      * Surface border AIMP
39      Mg.ECP.Larsson.14s7p.1s1p.0e-PBE-MgO(001) / AIMPLIB
40      pseudocharge
41      Mg9          2.1392500000    8.5570000000    2.1392500000  Angstrom
42      ...
43      Mg22         0.0000000000   -6.4177500000    2.1392500000  Angstrom
44      End of Basis
45      Basis set
46      * Bulk border AIMP
47      Mg.ECP.Larsson.14s7p.1s1p.0e-PBE-MgO(bulk) / AIMPLIB
48      pseudocharge
49      Mg23         2.1392500000    6.4177500000    0.0000000000  Angstrom
50      ...
51      Mg40         0.0000000000    0.0000000000   -4.2785000000  Angstrom
52      End of Basis
53      Basis set
54      * Surface AIMP
55      Mg.ECP.Larsson.0s.0s.0e-PBE-MgO(001) / AIMPLIB
56      pseudocharge
57      A1          -6.4177500000   -4.2785000000    2.1392500000  Angstrom
58      ...
59      A78          14.9747500000   12.8355000000    2.1392500000  Angstrom
60      End of Basis
61      Basis set
62      * Surface AIMP
63      O.ECP.Larsson.0s.0s.0e-PBE-MgO(001) / AIMPLIB
64      pseudocharge
65      A79          -4.2785000000   -4.2785000000    2.1392500000  Angstrom
66      ...
67      A164         14.9747500000   14.9747500000    2.1392500000  Angstrom
68      End of Basis
69      Basis set
70      * Bulk AIMP
71      Mg.ECP.Larsson.0s.0s.0e-PBE-MgO(001) / AIMPLIB
72      pseudocharge
73      A165         -4.2785000000    2.1392500000   -2.1392500000  Angstrom
74      ...
75      A830         14.9747500000   14.9747500000   -12.8355000000  Angstrom
76      End of Basis

```

```

77  Basis set
78  * Bulk AIMP
79  O.ECP.Larsson.0s.0s.0e-PBE-MgO(001) / AIMPLIB
80  pseudocharge
81  A831      -2.1392500000      2.1392500000      -2.1392500000  Angstrom
82  ...
83  A1508      14.9747500000      14.9747500000      -10.6962500000  Angstrom
84  End of Basis
85  xfield = $CurrDir/$Project.xfield
86
87  &SEWARD
88
89  &DMRGSCF
90  ActiveSpaceOptimizer = QCMAQUIS
91  Fiedler = on
92  DMRGSettings
93  nsweeps = 16
94  max_bond_dimension = 1000
95  donotdelete = 1
96  EndDMRGSettings
97  OptimizationSettings
98  fileorb = $CurrDir/$Project.dmrscf.h5
99  spin = 1
100  nactel = 22
101  EndOptimizationSettings

```

## References

- (S1) Dobrautz, W.; Weser, O.; Bogdanov, N. A.; Alavi, A.; Li Manni, G. Spin-Pure Stochastic-CASSCF via GUGA-FCIQMC Applied to Iron–Sulfur Clusters. *J. Chem. Theory Comput.* **2021**, *17*, 5684–5703, DOI: 10.1021/acs.jctc.1c00589.
- (S2) Balasubramani, S. G.; Chen, G. P.; Coriani, S.; Diedenhofen, M.; Frank, M. S.; Franzke, Y. J.; Furche, F.; Grotjahn, R.; Harding, M. E.; Hättig, C.; Hellweg, A.; Helmich-Paris, B.; Holzer, C.; Huniar, U.; Kaupp, M.; Khah, A. M.; Khani, S. K.; Müller, T.; Mack, F.; Nguyen, B. D.; Parker, S. M.; Perlt, E.; Rappoport, D.; Reiter, K.; Roy, S.; Rückert, M.; Schmitz, G.; Sierka, M.; Tapavicza, E.; Tew, D. P.; van Wüllen, C.; Voora, V. K.; Weigend, F.; Wodyński, A.; Yu, J. M. TURBO-MOLE: Modular program suite for *ab initio* quantum-chemical and condensed-matter simulations. *J. Chem. Phys.* **2020**, *152*, 184107, DOI: 10.1063/5.0004635.
- (S3) Cave, R. J.; Newton, M. D. Generalization of the Mulliken-Hush treatment for the calculation of electron transfer matrix elements. *Chem. Phys. Lett.* **1996**, *249*, 15–19, DOI: 10.1016/0009-2614(95)01310-5.
- (S4) Lévêque, J.; Rebolini, E.; Saúl, A.; Lepetit, M.-B. Theoretical study of the magnetic properties of BaNiF<sub>4</sub>. *Eur. Phys. J. B* **2021**, *94*, 214, DOI: 10.1140/epjb/s10051-021-00225-5.
- (S5) Lévêque, J.; Rebolini, E.; Lepetit, M.-B.; Saúl, A. Theoretical study of the magnetic properties of the CoCu<sub>2</sub>O<sub>3</sub> compound. *Phys. Rev. B* **2022**, *106*, DOI: 10.1103/physrevb.106.224402.
- (S6) Gellé, A.; Varignon, J.; Lepetit, M.-B. Accurate evaluation of magnetic coupling between atoms with numerous open shells: An *ab initio* method. *Europhys. Lett.* **2009**, *88*, 37003, DOI: 10.1209/0295-5075/88/37003.
- (S7) Kaufmann, K.; Baumeister, W.; Jungen, M. Universal Gaussian basis sets for an optimum representation of Rydberg and continuum wavefunctions. *J. Phys. B: At., Mol. Opt. Phys.* **1989**, *22*, 2223–2240, DOI: 10.1088/0953-4075/22/14/007.
- (S8) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **1980**, *48*, 157–173, DOI: 10.1016/0301-0104(80)80045-0.
- (S9) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. Main Group Atoms and Dimers Studied with a New Relativistic ANO Basis Set. *J. Phys. Chem. A* **2004**, *108*, 2851–2858, DOI: 10.1021/jp031064+.
- (S10) Dittrich, B.; Fabbiani, F. P. A.; Henn, J.; Schmidt, M. U.; Macchi, P.; Meindl, K.; Spackman, M. A. Azulene revisited: solid-state structure, invariom modeling and lattice-energy minimization of a classical example of disorder. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2018**, *74*, 416–426, DOI: 10.1107/s2052520618010120.
- (S11) Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. The multi-state CASPT2 method. *Chem. Phys. Lett.* **1998**, *288*, 299–306, DOI: 10.1016/s0009-2614(98)00252-8.

- (S12) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chem. Phys. Lett.* **2004**, *396*, 142–149, DOI: 10.1016/j.cplett.2004.08.032.
- (S13) Forsberg, N.; Malmqvist, P.-Å. Multiconfiguration perturbation theory with imaginary level shift. *Chem. Phys. Lett.* **1997**, *274*, 196–204, DOI: 10.1016/s0009-2614(97)00669-6.
- (S14) Aquilante, F.; Pedersen, T. B.; Lindh, R. Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals. *J. Chem. Phys.* **2007**, *126*, 194106, DOI: 10.1063/1.2736701.
- (S15) Aquilante, F.; Pedersen, T. B.; Lindh, R.; Roos, B. O.; de Merás, A. S.; Koch, H. Accurate *ab initio* density fitting for multiconfigurational self-consistent field methods. *J. Chem. Phys.* **2008**, *129*, 024113, DOI: 10.1063/1.2953696.
- (S16) Aquilante, F.; Malmqvist, P.-Å.; Pedersen, T. B.; Ghosh, A.; Roos, B. O. Cholesky Decomposition-Based Multiconfiguration Second-Order Perturbation Theory (CD-CASPT2): Application to the Spin-State Energetics of Co<sup>III</sup>(diiminato)(NPh). *J. Chem. Theory Comput.* **2008**, *4*, 694–702, DOI: 10.1021/ct700263h.
- (S17) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211, DOI: 10.1139/p80-159.
- (S18) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789, DOI: 10.1103/physrevb.37.785.
- (S19) Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377, DOI: 10.1063/1.464304.
- (S20) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652, DOI: 10.1063/1.464913.
- (S21) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297, DOI: 10.1039/b508541a.
- (S22) Titov, E. On the Low-Lying Electronically Excited States of Azobenzene Dimers: Transition Density Matrix Analysis. *Molecules* **2021**, *26*, 4245, DOI: 10.3390/molecules26144245.
- (S23) Malmqvist, P. Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO<sub>2</sub> and Cu<sub>2</sub>O<sub>2</sub> systems. *J. Chem. Phys.* **2008**, *128*, 204109, DOI: 10.1063/1.2920188.
- (S24) Sauri, V.; Serrano-Andrés, L.; Shahi, A. R. M.; Gagliardi, L.; Vancoillie, S.; Pierloot, K. Multiconfigurational Second-Order Perturbation Theory Restricted Active Space (RASPT2) Method for Electronic Excited States: A Benchmark Study. *J. Chem. Theory Comput.* **2011**, *7*, 153–168, DOI: 10.1021/ct100478d.

- (S25) Fujii, T.; Kashida, H.; Asanuma, H. Analysis of Coherent Heteroclustering of Different Dyes by Use of Threoninol Nucleotides for Comparison with the Molecular Exciton Theory. *Chem. – Eur. J.* **2009**, *15*, 10092–10102, DOI: 10.1002/chem.200900962.
- (S26) Kaiser, A.; Daoud, R. E.; Aquilante, F.; Kühn, O.; De Vico, L.; Bokarev, S. I. A Multiconfigurational Wave Function Implementation of the Frenkel Exciton Model for Molecular Aggregates. *J. Chem. Theory Comput.* **2023**, DOI: 10.1021/acs.jctc.3c00185, In press.
- (S27) Tenorio, B. N. C.; Decleva, P.; Coriani, S. Multi-reference approach to the computation of double core-hole spectra. *J. Chem. Phys.* **2021**, *155*, 131101, DOI: 10.1063/5.0062130.
- (S28) Tenorio, B. N. C.; Voß, T. A.; Bokarev, S. I.; Decleva, P.; Coriani, S. Multireference Approach to Normal and Resonant Auger Spectra Based on the One-Center Approximation. *J. Chem. Theory Comput.* **2022**, *18*, 4387–4407, DOI: 10.1021/acs.jctc.2c00252.
- (S29) Ferté, A.; Palaudoux, J.; Penent, F.; Iwayama, H.; Shigemasa, E.; Hikosaka, Y.; Soejima, K.; Ito, K.; Lablanquie, P.; Taïeb, R.; Carniato, S. Advanced Computation Method for Double Core Hole Spectra: Insight into the Nature of Intense Shake-up Satellites. *J. Phys. Chem. Lett.* **2020**, *11*, 4359–4366, DOI: 10.1021/acs.jpcllett.0c01167.
- (S30) Eberhardt, W.; Stöhr, J.; Feldhaus, J.; Plummer, E. W.; Sette, F. Correlation between Electron Emission and Fragmentation into Ions following Soft-X-Ray Excitation of the N<sub>2</sub> Molecule. *Phys. Rev. Lett.* **1983**, *51*, 2370–2373, DOI: 10.1103/physrevlett.51.2370.
- (S31) Kaliakin, D. S. OpenMolcas–GAMESS/Tinker interface repository. <https://gitlab.com/interface-between/openmolcas-and-gamess-us-with-tinker>.
- (S32) Ferré, N.; Ángyán, J. G. Approximate electrostatic interaction operator for QM/MM calculations. *Chem. Phys. Lett.* **2002**, *356*, 331–339, DOI: 10.1016/s0009-2614(02)00343-3.
- (S33) Aquilante, F.; Autschbach, J.; Baiardi, A.; Battaglia, S.; Borin, V. A.; Chibotaru, L. F.; Conti, I.; De Vico, L.; Delcey, M.; Fdez. Galván, I.; Ferré, N.; Freitag, L.; Garavelli, M.; Gong, X.; Knecht, S.; Larsson, E. D.; Lindh, R.; Lundberg, M.; Malmqvist, P. Å.; Nenov, A.; Norell, J.; Odelius, M.; Olivucci, M.; Pedersen, T. B.; Pedraza-González, L.; Phung, Q. M.; Pierloot, K.; Reiher, M.; Schapiro, I.; Segarra-Martí, J.; Segatta, F.; Seijo, L.; Sen, S.; Sergentu, D.-C.; Stein, C. J.; Ungur, L.; Vacher, M.; Valentini, A.; Veryazov, V. Modern quantum chemistry with [Open]Molcas. *J. Chem. Phys.* **2020**, *152*, 214117, DOI: 10.1063/5.0004835.
- (S34) Huix-Rotllant, M.; Ferré, N. Analytic Energy, Gradient, and Hessian of Electrostatic Embedding QM/MM Based on Electrostatic Potential-Fitted Atomic Charges Scaling Linearly with the MM Subsystem Size. *J. Chem. Theory Comput.* **2020**, *17*, 538–548, DOI: 10.1021/acs.jctc.0c01075.

- (S35) Yang, X.; Manathunga, M.; Gozem, S.; Léonard, J.; Andruniów, T.; Olivucci, M. Quantum–classical simulations of rhodopsin reveal excited-state population splitting and its effects on quantum efficiency. *Nat. Chem.* **2022**, *14*, 441–449, DOI: 10.1038/s41557-022-00892-6.
- (S36) Pedraza-González, L.; Barneschi, L.; Padula, D.; De Vico, L.; Olivucci, M. Evolution of the Automatic Rhodopsin Modeling (ARM) Protocol. *Top Curr Chem (Z)* **2022**, *380*, DOI: 10.1007/s41061-022-00374-w.
- (S37) Pedraza-González, L.; De Vico, L.; del Carmen Marín, M.; Fanelli, F.; Olivucci, M. *a*-ARM: Automatic Rhodopsin Modeling with Chromophore Cavity Generation, Ionization State Selection, and External Counterion Placement. *J. Chem. Theory Comput.* **2019**, *15*, 3134–3152, DOI: 10.1021/acs.jctc.9b00061.
- (S38) Melaccio, F.; del Carmen Marín, M.; Valentini, A.; Montisci, F.; Rinaldi, S.; Cherubini, M.; Yang, X.; Kato, Y.; Stenrup, M.; Orozco-Gonzalez, Y.; Ferré, N.; Luk, H. L.; Kandori, H.; Olivucci, M. Toward Automatic Rhodopsin Modeling as a Tool for High-Throughput Computational Photobiology. *J. Chem. Theory Comput.* **2016**, *12*, 6020–6034, DOI: 10.1021/acs.jctc.6b00367.
- (S39) Paolino, M.; Giovannini, T.; Manathunga, M.; Latterini, L.; Zampini, G.; Pieron, R.; Léonard, J.; Fusi, S.; Giorgi, G.; Giuliani, G.; Cappelli, A.; Cappelli, C.; Olivucci, M. On the Transition from a Biomimetic Molecular Switch to a Rotary Molecular Motor. *J. Phys. Chem. Lett.* **2021**, *12*, 3875–3884, DOI: 10.1021/acs.jpcllett.1c00526.
- (S40) Chatterjee, G.; Jha, A.; Blanco-Gonzalez, A.; Tiwari, V.; Manathunga, M.; Duan, H.-G.; Tellkamp, F.; Prokhorenko, V. I.; Ferré, N.; Dasgupta, J.; Olivucci, M.; Miller, R. J. D. Torsionally broken symmetry assists infrared excitation of biomimetic charge-coupled nuclear motions in the electronic ground state. *Chem. Sci.* **2022**, *13*, 9392–9400, DOI: 10.1039/d2sc02133a.
- (S41) Filatov, M. Spin-restricted ensemble-referenced Kohn–Sham method: basic principles and application to strongly correlated ground and excited states of molecules. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2015**, *5*, 146–167, DOI: 10.1002/wcms.1209.
- (S42) Ferré, N., Filatov, M., Huix-Rotllant, M., Eds. *Density-Functional Methods for Excited States*; Topics in Current Chemistry 368; Springer International Publishing, 2016; DOI: 10.1007/978-3-319-22081-9.
- (S43) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639–9646, DOI: 10.1063/1.473863.
- (S44) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N<sub>2</sub>, and H<sub>2</sub>O. *Chem. Phys. Lett.* **1998**, *286*, 243–252, DOI: 10.1016/s0009-2614(98)00111-0.
- (S45) Schwenke, D. W. The extrapolation of one-electron basis sets in electronic structure calculations: How it should work and how it can be made to work. *J. Chem. Phys.* **2005**, *122*, 014107, DOI: 10.1063/1.1824880.

- (S46) Neese, F.; Valeev, E. F. Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated *ab initio* Methods? *J. Chem. Theory Comput.* **2011**, *7*, 33–43, DOI: 10.1021/ct100396y.
- (S47) Granatier, J. Extrapolation of Atomic Natural Orbitals of basis set to complete basis set limit. *Acta Chim. Slovaca* **2017**, *10*, 159–164, DOI: 10.1515/acs-2017-0026.
- (S48) Widmark, P.-O.; Persson, B. J.; Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theor. Chim. Acta* **1991**, *79*, 419–432, DOI: 10.1007/bf01112569.
- (S49) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. New Relativistic ANO Basis Sets for Transition Metal Atoms. *J. Phys. Chem. A* **2005**, *109*, 6575–6579, DOI: 10.1021/jp0581126.
- (S50) Zobel, J. P.; Widmark, P.-O.; Veryazov, V. The ANO-R Basis Set. *J. Chem. Theory Comput.* **2020**, *16*, 278–294, DOI: 10.1021/acs.jctc.9b00873.
- (S51) Kelleher, D. E.; Martin, W. C.; Wiese, W. L.; Sugar, J.; Fuhr, J. R.; Olsen, K.; Musgrove, A.; Mohr, P. J.; Reader, J.; Dalton, G. R. The New NIST Atomic Spectra Database. *Phys. Scr.* **1999**, *T83*, 158, DOI: 10.1238/physica.topical.083a00158.
- (S52) Kantorovich, L. N.; Gillan, M. J.; White, J. A. Adsorption of atomic oxygen on the MgO (100) surface. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 2075, DOI: 10.1039/ft9969202075.
- (S53) Larsson, E. D.; Krośnicki, M.; Veryazov, V. A program system for self-consistent embedded potentials for ionic crystals. *Chem. Phys.* **2022**, *562*, 111549, DOI: 10.1016/j.chemphys.2022.111549.