ONETEP to GENNBO FILE.47 Input Parameters

Louis Lee
TCM Group, Cavendish Laboratory
19 JJ Thomson Ave, Cambridge CB3 0HE, UK
lpl24@cam.ac.uk

June 30, 2012

1 Standalone NBO 5 Program GENNBO

The standalone version of the NBO program (GENNBO)[1] accepts parameters from an input ASCII free-format FILE.47 containing atomic coordinates and matrix information, as printed by ONETEP if the Natural Populataion Analysis[2] subroutine is called during a PROPERTIES calculation, by specifying the keyword write_nbo: T. The NBO formalism allows one transform a converged 1-particle wavefunction in an atom-centred bases into a set of highly-local 'Natural Bond Orbitals', which are one and two (or three)-centred 'lone' and 'bond' pairs recognizable as chemical bonds from a classical Lewis structure standpoint.[3] Details of the NBO formalism are discussed elsewhere.[2–4]

1.1 Compiling GENNBO

Compilation of the standalone GENNBO does not involve ONETEP in any way. As of writing, the latest NBO release is version 5.9.[1] Compilation instructions are listed here for convenience, based on some trial-and-error when the arcane g77 compiler listed in the NBO manual is unavailable.

To compile GENNBO, first, compile the activator, enable.f:

```
gfortran -o enable enable.f
```

then run the enable program. Complete the selections to generate the standalone GENNBO source gennbo.f.

By default, GENNBO limits the number of atoms and basis in the FILE.47 input to 200 and 2000 respectively. This can be increased by replacing all instances of MAXATM = 200 and MAXBAS = 2000 to a user-specified value, up to a limit of 999 and 9999 respectively (higher values are possible, albeit accompanied by illegible output due to format overflow. In principle one could modify the code even further to remedy this issue.).

The following command should compile GENNBO correctly on x64 architectures, when no modification is made to the gennbo.f source:

```
gfortran -fdefault-integer-8 -fno-sign-zero -m64 -o gennbo gennbo.f
ifort -i8 -m64 -f77rtl -o gennbo gennbo.f
```

For the 32-bit version, integer length should be set to 4 bytes instead (e.g. -i4 in ifort). If MAXATM and MAXBAS have been increased then the memory model should also be set to allow data > 2 GB, by adding a -mcmodel=medium flag. For ifort, an additional -shared-intel flag is most likely necessary.

Then, to run:

```
gennbo < FILE.47 > output.out
```

2 ONETEP NPA Generation Routine

The Natural Population Analysis[2] method of computing atomic charges is implemented in ONETEP. The routine transforms the set of non-orthogonal, optimized NGWFs into a set of orthogonal atom-centred 'Natural Atomic Orbitals' (NAOs) via an 'occupancy-weighted symmetric orthogonalization' procedure, which serves to maximise the resemblance of the final orthogonal orbitals to their initial non-orthogonal parents (a la Löwdin orthogonalization), weighted according to the parent orbital occupancies. Therefore, vacant, highly-diffuse orbitals are free to distort to achieve orthogonality with their more highly-preserved occupied counterpart. This ensures that the final NAO population (the 'Natural Population') remains stable with respect to basis set size.

Once in the NAO basis, further transformations such as pair-block density matrix diagonalization produce the final set of NBOs – these procedures are performed by NBO 5 from the `FILE.47` output of ONETEP, which contains relevant matrices in the NAO basis. The NAO routine is performed internally in ONETEP as NBO 5 requires pseudo-atomic orbitals (such as Gaussian-type orbitals) with free-atom symmetries and orthogonality within each atom, a property not rigorously satisfied by the optimized NGWFs.

The NPA module in ONETEP performs at its best for large systems when compiles with the ScaLapack linear algebra package, as it takes advantage of the distributed memory storage of dense global matrices, such as the inverse square root of the overlap matrix that needs to be computed for the 'occupancy-weighted symmetric orthogonalization' step. This has the unfortunate side effect of rendering the NAO transformation a cubic-scaling method. However, this step occurs only once during the routine, and should be comparable to the time needed to generate canonical molecular orbitals.

3 List of Available Parameters

NAO Generation (Default)				
Keyword	Type	Default	Level	Description
write_nbo	L	F	B	Enables Natural Population Analysis (NPA) and writing of GENNBO input file <seedname>_nao_nbo.47
nbo_init_lclowdin	L	T	E	Performs atom-local Löwdin orthogonalisation on NGWFs as the first step before constructing NAOs
nbo_write_lclowdin	L	F	E	Writes full matrices (all atoms) in the atom-local Löwdin-orthogonalized basis to FILE.47 (For reference/testing/comparison purposes). Output will be <seedname>_lclowdin_nbo.47
nbo_write_npacomp	L	F	B	Writes NAO charges for all orbitals to standard output
nbo_write_dipole	L	F	B	Computes and writes dipole matix to FILE.47
nbo_scale_dm	L	T	E	Scales partial density matrix output to <seedname>_nao_nbo.47 in order to achieve charge integrality
nbo_scale_spin	L	T	E	Scales α and β spins independently to integral chrage when partial matrices are printed and nbo_scale_dm = T. Inevitably means spin density values from GENNBO are invalid and one should calculate them manually from the α and β NPA populations.
nbo_write_species	B	N/A	B	Block of lists of species to be included in the partial matrix output of <seedname>_nao_nbo.47. If not present all atoms will be included. E.g. specified will default to AUTO. E.g.: %block nbo_write_species C1 H1 %endblock nbo_write_species
nbo_species_ngwflabel	B	AUTO	I	Optional user-defined (false) lm-label for NGWFs according to GENNBO convention. Species not specified will default to AUTO. E.g.: %block nbo_species_ngwflabel C1 "1N 151N 152N 153N" H1 "AUTO" %endblock nbo_species_ngwflabel -N suffix denotes NMB orbital. If 'SOLVE' orbitals are used, this block should be present as 'AUTO' initialisation assumes orbitals were also initialised as 'AUTO'.
nbo_aopnao_scheme	T	ORIGINAL	E	The AO to PNAO scheme to use. Affects the 'lm-averaging' and diagonalisation steps in the initial AO to PNAO, NRB lm-averaging, and rediagonaliation transformations (the '**N**' transformations in [2]). For testing purposes only - so far none of the other schemes apart from ORIGINAL works. Possbile values are: ORIGINAL - default, as in [2] with lm-averaging DIAGONALIZATION - Diagonalises entire atom-centred sub-block w/o lm-averaging or splitting between different angular channels. NONE - Skips all '**N**' transformations.
nbo_pnao_analysis	L	F	E	Perform s/p/d/f analysis on the PNAOs (analogous to ngwf_analysis)

Orbital Plotting				
Keyword	Type	Default	Level	Description
plot_nbo	L	F	B	Instructs ONETEP to read the relevant orbital transformation output from GENNBO, determined by the flag nbo_plot_orbtype and plots the orbitals specified in %block nbo_list_plotnbo. write_nbo and plot_nbo are mutually exclusive. Scalar field plotting must be enabled (e.g. cube_format = T).
nbo_plot_orbtype	T	N/A	B	The type of GENNBO-generated orbitals to read and plot. Possible values and their associated GENNBO transformation files must be present, as follows: NAO - <seedname>_nao.33 NHO - <seedname>_nao.35 NBO - <seedname>_nao.37 NLMO - <seedname>_nao.39 NLMO is only defined for the full system i.e. partitioned FILE.47 will give meaningless NLMOs. Except for NLMO, adding a 'P' prefix e.g. 'PNAO' to the value of nbo_plot_orbtype causes the non-orthogonalised PNAOs to be used in plotting instead of NAOs. PNAOs are of the normal type, i.e. when RPNAO = F in GENNBO (default).
nbo_list_plotnbo	B	N/A	B	The list of nbo_plot_orbtype orbitals to be plotted, identified by their indices as in the GENNBO output. Specify each index on a new line.

Output files:	
<seedname>_nao_nbo.47	Always written. Contains partial matrices according to %block nbo_write_species.
<seedname>_lclowdin_nbo.47	Written if nbo_write_lclowdin = T. Always contains all atoms in the atom-local Löwdin-orthogonalized basis. If nbo_init_lclowdin = T and all atoms are included <seedname>_lclowdin_nbo.47 = <seedname>_nao_nbo.47 except for ordering of atomic centers (will be fixed in newer releases).
<seedname>_nao_atomindex.dat	Contains mapping of atomic indices of the potentially subset of the full system in <seedname>_nao_nbo.47 to the real atomic index of the full system (since labels have to be consecutive). Real atomic index refers to original input order in ONETEP.
<seedname>_inittr_nao_nbo.dat	Raw NGWF to NAO transformation read for plotting (i.e. when plot_nbo = T).
<seedname>_inittr_pnao_nbo.dat	Raw NGWF to PNAO transformation read for plotting (i.e. when plot_nbo = T). PNAOs are of the 'normal' type, i.e when RPNAO = F in GENNBO.
<seedname>_nbo_DEBUG.dat	Contains various debugging info. Only written if compiled in debug mode.

4 Notes

4.1 Orbital labelling with pseudoatomic solver

- GENNBO labels in `%block nbo_species_ngwflabel` should always be explicitly given when `SOLVE` is used to initialise the NGWFs. The label string is however limited to 80 characters in ONETEP, which should be fine up to $1s2sp3spd4sp$. This will be fixed later unless it is urgently required.
- Make sure the orbitals selected for plotting are valid. The NPA routine assumes that the appropriate transformation file from GENNBO in the same directory is correct, and only complains if it encounters an EOF, but not if the wrong transformation file is given (e.g. from a different system with a larger basis).
- Do not rename the GENNBO-generated transformation files. ONETEP expects them to have the name `<seedname>_nao.xx`.

4.2 Orbital plotting

- In order to plot the various orbitals, first run the output `FILE.47` through GENNBO to obtain the relevant orbital vectors. Refer to the NBO 5 manual for details on how to print these (e.g. to print NBOs in the input `FILE.47` basis, set `AONBO=W` in the `$NBO` block).
- For some reason, the `PLOT` keyword itself in GENNBO doesn't work. This might have something to do with the '`ORTHO` bug'.

4.3 '`ORTHO` bug'

The NBO 5 program up till circa April/May 2011 had a bug whereby specifying the `ORTHO` flag causes the program to crash. The NBO 5 developers seem to have fixed most of this and given me the an updated version, but residual bug could remain (have they made the fix a general release yet?). This is of course fixable by running the `<seedname>_lclowdin_nbo.47` file through GENNBO instead, albeit this would mean one can't do DM partitioning.

5 Example Usage

5.1 Obtaining 2^{nd}-order Perturbation Estimates of the $n \rightarrow \sigma^*$ Secondary Hyperconjugation in Water Dimer (Hydrogen Bond)

The hydrogen bond stabilization in water dimer can be attributed to the non-classical 'charge transfer' interaction between two water molecules due to delocalization of the electronic charge from the oxygen lone pair n of the donor monomer to the σ^* O–H antibond of the acceptor.[3] The expansion of the variational space to included non-Lewis, formally vacant antibond NBOs leads to an energetic lowering compared to the ideal Lewis configuration (all Lewis NBO occupancy = 2 e), which can be estimated via 2^{nd}-order perturbation theory as the 'charge transfer' energetic component of the dimer interaction.

From a converged SCF calculation in ONETEP using the reference coordinates below (given in Angstroms):

```
%block  positions_abs
"ang"
 O  10.6080354926368 12.5000150953008 12.5705695516353
 H  10.4341376488693 12.5000119731552 13.5119746410112
 H1 11.5729802892758 12.5000098564464 12.5000098564464
 H  13.9638274701977 13.2691512541917 12.2000071259853
 O1 13.4760438789916 12.5000098564464 12.5000098564464
 H  13.9638258826660 11.7308667653340 12.2000083960106
%endblock  positions_abs
```

with the pseudoatomic solver employing a minimal NGWF basis (1 NGWF on H, 4 on O) with a 10.0 a_0 NGWF radius cutoff, PBE exchange-correlation functional, norm-conserving pseudopotential with pseudized $1s$ core for O, and a 1200 eV psinc cutoff in a 25.0 a_0 cubic simulation cell, one should run a `PROPERTIES` calculation with the additional keywords as such:

```
write_nbo: T
%block species_ngwflabel
 H  "1N"
 O  "1N 152N 153N 151N"
 H1 "1N"
 O1 "1N 152N 153N 151N"
%endblock species_ngwflabel
```

where the `species_ngwflabel` block tells the NPA routine in ONETEP how to label each NGWF. The order of m for each l in the $Y(l,m)$ isn't straightforward, and follows the pattern of e.g. "152 153 151" i.e. $m = \{-1, 0, 1\}$ for $l = 1$, and "251 253 255 252 254" for $l = 2$. I've yet to look at how others are arranged, though this is not very important unless one is interested in 'NHO Directionality and Bond Bending' analysis, as in the NBO scheme, all m of the same l are treated equally. The order of each $Y(l,m)$ should follow that of the pseudoatomic solver, which does them in principal quantum number (n) increments (with multiple-ζ basis, the split-valence set of $Y(l,m)$ probably comes first i.e. $Y^{\zeta 1}(l,m)$ then $Y^{\zeta 2}(l,m)$ before the next n. The "N" suffix denotes valence orbital in the ground state, which in the case of H, "1N" is the $1s$ orbital. Make sure the correct orbitals are marked as valence as they would appear in the ground state (even if the pseudoatomic solver basis was initialized in an excited configuration). In this example, the pseudoatomic solver block would have explicitly been:

```
%block species_atomic_set
 H  "SOLVE conf=1s1"
 O  "SOLVE conf=1sX 2s2 2p4"
%endblock species_atomic_set
```

ONETEP should run and produce an NPA output listing the NPA charges on each atom, and print a `<seedname>_nao_nbo.47` file. This `.47` file serves as the input for GENNBO.

If we wanted to generate NBOs and visualize them, insert the keyword `AONBO=W` in the `$NBO` block of the `.47` file before running it through GENNBO. GENNBO will output a report containing NBO information, including the 2$^{\text{nd}}$-order perturbation estimates, and a `.37` file containing the NBO vectors in terms of the `.37` input basis (don't change any of the `.47`, `.37` etc. filenames).

First, we can see that the 2$^{\text{nd}}$-order perturbation report shows one prominenet interaction, namely one between the occupied lone pair of oxygen from one H_2O unit (`LP ( 2) O 5`) to the O–H antibond of the other (`BD*( 2) O 1- H 3`) with an estimate of 15.32 kcal/mol, corresponding to the hydrogen bond in water dimer:

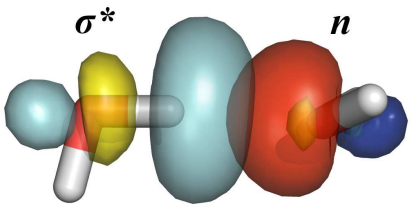
```
SECOND ORDER PERTURBATION THEORY ANALYSIS OF FOCK MATRIX IN NBO BASIS

    Threshold for printing:   0.50 kcal/mol
   (Intermolecular threshold: 0.05 kcal/mol)
                                                              E(2)  E(j)-E(i) F(i,j)
     Donor NBO (i)              Acceptor NBO (j)       kcal/mol   a.u.    a.u.
===============================================================================

within unit  1
      None above threshold

from unit  1 to unit  2
  2. BD ( 1) O 1- H 3       11. BD*( 1) H 4- O 5         0.08    0.67    0.007
  2. BD ( 1) O 1- H 3       12. BD*( 1) O 5- H 6         0.08    0.67    0.007

from unit  2 to unit  1
  3. BD ( 1) H 4- O 5       10. BD*( 1) O 1- H 3         0.10    0.83    0.008
  4. BD ( 1) O 5- H 6       10. BD*( 1) O 1- H 3         0.10    0.83    0.008
  7. LP ( 1) O 5            10. BD*( 1) O 1- H 3         0.18    0.49    0.008
  8. LP ( 2) O 5            10. BD*( 1) O 1- H 3        15.32    0.60    0.085

within unit  2
      None above threshold
```

Noting down the orbital numbers, we can then proceed to plot them by running another properties calculation in ONETEP with the following block:

```
write_nbo   : F
plot_nbo    : T
cube_format : T
nbo_plot_orbtype : NBO
%block nbo_list_plotnbo
  8
 10
%endblock nbo_list_plotnbo
```

where `write_nbo` needs to be set to `F`. ONETEP will then read the `<seedname>_inittr_nao_nbo.dat` file printed during the first run and the `.37` file to plot the orbitals specified in the `nbo_list_plotnbo` block into Gaussian cube files. Our example result is displayed in Figure 1.

Figure 1: Example plots of the $n \rightarrow \sigma^*$ hyperconjugation in water dimer. Oxygen lone pair n in red/blue (+/-), O–H antibond σ^* in cyan/yellow (+/-). Isosurface contour value at 0.075 a.u.



5.2 Notes on Selectively Passing sub-region sub-matrices into GENNBO

To circumvent the limitations on system size in GENNBO, and for convenience, we could output only matrix elements corresponding to atoms within a selected sub-region of a large system. To do so, during an NPA analysis run (not plotting) within a properties run in ONETEP, the following should be specified:

```
%block nbo_write_species
O1
H1
C1
⋮
%endblock nbo_write_species
```

ONETEP would then print only matrix elements belongning to species specified by the labels in the `%block nbo_write_species` block to `<seedname>_nao_nbo.47`. Due to GENNBO insisting on integral charges, the density matrix in the `.47` file is re-scaled downwards to the nearest lowest integral number, to avoid the possibility of orbitals having occupancies > 2 e, which also annoys GENNBO. To minimize the impact of this technical re-scaling to the NBO results, a sufficiently-sized partition should be chosen in `%block nbo_write_species` so that $1/N_e << 1$, where N_e is the number of electrons in the partition.

The final results fron NBO analysis that depend on the density matrix will then need to be de-scaled to arrive at the correct value (e.g. NPA charges, NBO occupancies, 2^{nd}-order perturbation estimates, while orbital energies don't require de-scaling).

Note that the region included in `%block nbo_write_species` should have buffer atoms, which minimally should include the next-nearest neighbour atom bonded to the last atom in the selection – that way, the severing of a bond would only affect NBOs centred on the buffer atom, and not anywhere else.

As a final note, there is a possibility that during an NBO search, slightly different NBO pictures are obtained when passing only part of the matrix as compared to analyzing the full system – this can be caused by the fact that during an NBO search, the NBO 5 program iterates through different occupancy thresholds (n_{min}) for deciding upon whether an orbital is a lone pair/NBO. If one is padentic about this, then n_{min} can be fixed by specifying the `THRESH =` n_{min} keyword manually in the `$NBO` block in the `.47` file, where n_{min} is defined by the user.

References

[1] E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, F. Weinhold; NBO 5.9 (http://www.chem.wisc.edu/~nbo5) & the NBO 5.9 Manual, Theoretical Chemistry Institute, University of Wisconsin, Madison, WI.

[2] A. E. Reed, R. B. Weinstock, F. Weinhold *J. Chem. Phys.* **1985,** *83,* 735-746.

[3] A. E. Reed, L. A. Curtiss, F. Weinhold *Chem. Rev.* **1988,** *88,* 899-926.

[4] A. D. MacKerell, Jr., B. Brooks, C. L. Brooks III, L. Nilsson, B. Roux, Y. Won, M. Karplus, in Encyclopedia of Computational Chemistry; R. Schleyer et al. Eds.; John Wiley & Sons, Chichester, **1998**; Vol. 3, Chapter 'Natural Bond Orbital Methods', pp 1792-1811.