The Jastrow Factor

Pablo López Ríos
26-Jul-2009

The Jastrow Factor

- Typical trial wave function – Slater-Jastrow:

$$\Psi_T(\boldsymbol{R}) = e^{J(\boldsymbol{R})}\Psi_S(\boldsymbol{R})$$

- In VMC, Jastrow describes correlations
- DMC is like VMC with perfect Jastrow; actual Jastrow merely stabilizes method
- VMC is much simpler/cheaper than DMC – pushing VMC quality towards DMC is good

The Jastrow Factor

- Jastrow function: sum of terms, e.g.,

$$J(\boldsymbol{R}) = J_{e-e}(\boldsymbol{R}) + J_{e-n}(\boldsymbol{R}) + J_{e-e-n}(\boldsymbol{R}) + \ldots$$

- Term: sum over particle groups of a function, e.g.,

$$J_{e-e-n}(\boldsymbol{R}) = \sum_{i,j}^{N} \sum_{I}^{M} \Omega(\boldsymbol{r}_{ij}, \boldsymbol{r}_{iI}, \boldsymbol{r}_{jI})$$

- Function: expansion on basis of functions of single relative position vector, e.g.

$$\Omega(\boldsymbol{r}_{ij}, \boldsymbol{r}_{iI}, \boldsymbol{r}_{jI}) = \sum_{\nu_{ij}}^{p} \sum_{\mu_{iI}, \mu_{jI}}^{q} \lambda_{\nu_{ij}, \mu_{iI}, \mu_{jI}} \phi_{\nu_{ij}}(\boldsymbol{r}_{ij}) \theta_{\mu_{iI}}(\boldsymbol{r}_{iI}) \theta_{\mu_{jI}}(\boldsymbol{r}_{jI})$$

Why Generalize?

- All Jastrow terms constructed the same way
- No need to implement **new terms** (e.g., four-body)
- Easy to implement **new functional bases**
- Easy to add **anisotropies**
- Easy to add dependencies on **external potentials**

General Jastrow Term: definition

$$J_{n,m}(\boldsymbol{R})=\frac{1}{n!\,m!}\sum_{i_1\neq\ldots\neq i_n}^{N}\ \sum_{I_{\backslash}\neq\ldots\neq I_m}^{M}\ \sum_{\{\nu_{i_\alpha i_\beta}\}_{\alpha<\beta}^{n}}^{p}\ \sum_{\{\mu_{i_\alpha I_\gamma}\}_{\alpha,\gamma}^{n,m}}^{q}\lambda_{\{\nu\}\{\mu\}}^{\{P\}\{S\}}\times$$

$$\times\prod_{\alpha<\beta}^{n}\Phi_{\nu_{i_\alpha i_\beta}}^{P_{\sigma(i_\alpha)\sigma(i_\beta)}}\left(\boldsymbol{r}_{i_\alpha i_\beta}\right)\prod_{\alpha,\gamma}^{n,m}\Theta_{\mu_{i_\alpha I_\gamma}}^{S_{\sigma(i_\alpha)I_\gamma}}\left(\boldsymbol{r}_{i_\alpha I_\gamma}\right)$$

Definition

$$J_{n,m}(\boldsymbol{R})=\sum_{i_1<\ldots<i_n}^{N}\ \sum_{I_1<\ldots<I_m}^{M}\ \sum_{\{\nu_{i_\alpha i_\beta}\}_{\alpha<\beta}^{n}}^{p}\ \sum_{\{\mu_{i_\alpha I_\gamma}\}_{\alpha,\gamma}^{n,m}}^{q}\lambda_{\{\nu\}\{\mu\}}^{\{P\}\{S\}}\times$$

$$\times\prod_{\alpha<\beta}^{n}\Phi_{\nu_{i_\alpha i_\beta}}^{P_{\sigma(i_\alpha)\sigma(i_\beta)}}\left(\boldsymbol{r}_{i_\alpha i_\beta}\right)\prod_{\alpha,\gamma}^{n,m}\Theta_{\mu_{i_\alpha I_\gamma}}^{S_{\sigma(i_\alpha)I_\gamma}}\left(\boldsymbol{r}_{i_\alpha I_\gamma}\right)$$

Simplified form

“One electron” Jastrow

$$J^{i}_{n,m}(\boldsymbol{R})=\sum_{\substack{i_2<\ldots<i_n\\ i_1=i,\,i_\alpha\neq i}}^{N}\ \sum_{I_1<\ldots<I_m}^{M}\ \sum_{\{\nu_{i_\alpha i_\beta}\}_{\alpha<\beta}^{n}}^{p}\ \sum_{\{\mu_{i_\alpha I_\gamma}\}_{\alpha,\gamma}^{n,m}}^{q}\lambda^{\{P\}\{S\}}_{\{\nu\}\{\mu\}}\times$$

$$\times\prod_{\alpha<\beta}^{n}\Phi^{P_{\sigma(i_\alpha)\sigma(i_\beta)}}_{\nu_{i_\alpha i_\beta}}\left(\boldsymbol{r}_{i_\alpha i_\beta}\right)\prod_{\alpha,\gamma}^{n,m}\Theta^{S_{\sigma(i_\alpha)I_\gamma}}_{\mu_{i_\alpha I_\gamma}}\left(\boldsymbol{r}_{i_\alpha I_\gamma}\right)$$

One-electron Jastrow for i-th particle

Properties:

$$J_{n,m}(\boldsymbol{R}')-J_{n,m}(\boldsymbol{R})=J^{i}_{n,m}(\boldsymbol{R}')-J^{i}_{n,m}(\boldsymbol{R})$$

1. If two configurations only differ in the position of particle i, change in total Jastrow is change in one-electron Jastrow

$$J_{n,m}=\sum_{i=1}^{N}\frac{1}{n}J^{i}_{n,m}$$

2. Total Jastrow value can be computed from one-electron Jastrow for all electrons

Gradient

$$\nabla_i J_{n,m}(\boldsymbol{R}) = \sum_{\substack{i_2<\ldots<i_n \\ i_1=i,\, i_\alpha\neq i}}^{N} \sum_{I_1<\ldots<I_m}^{M} \sum_{\{\nu_{i_\alpha i_\beta}\}_{\alpha<\beta}^{n}}^{p} \sum_{\{\mu_{i_\alpha I_\gamma}\}_{\alpha,\gamma}^{n,m}}^{q} \times$$

$$\times \lambda_{\{\nu\}\{\mu\}}^{\{P\}\{S\}} \prod_{\alpha<\beta}^{n} \Phi_{\nu_{i_\alpha i_\beta}}^{P_{\sigma(i_\alpha)\sigma(i_\beta)}}\left(\boldsymbol{r}_{i_\alpha i_\beta}\right) \prod_{\alpha,\gamma}^{n,m} \Theta_{\mu_{i_\alpha I_\gamma}}^{S_{\sigma(i_\alpha) I_\gamma}}\left(\boldsymbol{r}_{i_\alpha I_\gamma}\right) \times$$

$$\times \left(\sum_{\beta=2}^{n} \frac{\nabla_i \Phi_{\nu_{ii_\beta}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right)}{\Phi_{\nu_{ii_\beta}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right)} + \sum_{\gamma=1}^{m} \frac{\nabla_i \Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i) I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right)}{\Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i) I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right)} \right)$$

In this form, the gradient is calculated simply by multiplying the running accumulator for the value times a simple sum of gradients of basis functions.

Basis functions being exactly zero are an issue, but I don't expect this to ever happen in practice. The code currently ignores contributions to the gradient from zero-valued functions.

Laplacian

$$\nabla_i^2 J_{n,m}(\boldsymbol{R}) = \sum_{\substack{i_2<\ldots<i_n \\ i_1=i,\, i_\alpha \neq i}}^{N} \; \sum_{I_1<\ldots<I_m}^{M} \; \sum_{\{\nu_{i_\alpha i_\beta}\}_{\alpha<\beta}^{n}}^{p} \; \sum_{\{\mu_{i_\alpha I_\gamma}\}_{\alpha,\gamma}^{n,m}}^{q} \lambda_{\{\nu\}\{\mu\}}^{\{P\}\{S\}} \times$$

$$\times \prod_{\alpha<\beta}^{n} \Phi_{\nu_{i_\alpha i_\beta}}^{P_{\sigma(i_\alpha)\sigma(i_\beta)}}\left(\boldsymbol{r}_{i_\alpha i_\beta}\right) \prod_{\alpha,\gamma}^{n,m} \Theta_{\mu_{i_\alpha I_\gamma}}^{S_{\sigma(i_\alpha) i_\gamma}}\left(\boldsymbol{r}_{i_\alpha I_\gamma}\right) \times$$

$$\times \left[\left(\sum_{\beta=2}^{n} \frac{\nabla_i \Phi_{\nu_{ii_\beta}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right)}{\Phi_{\nu_{ii_\beta}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right)} + \sum_{\gamma=1}^{m} \frac{\nabla_i \Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right)}{\Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right)} \right)^2 + \right.$$

$$+ \sum_{\beta=2}^{n} \left(\frac{\nabla_i^2 \Phi_{\nu_{ii_\beta}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right)}{\Phi_{\nu_{iI(\beta)}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right)} - \left(\frac{\nabla_i \Phi_{\nu_{ii_\beta}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right)}{\Phi_{\nu_{ii_\beta}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right)} \right)^2 \right) +$$

$$\left. + \sum_{\gamma=2}^{m} \left(\frac{\nabla_i^2 \Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right)}{\Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right)} - \left(\frac{\nabla_i \Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right)}{\Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right)} \right)^2 \right) \right]$$

Likewise for
the Laplacian

The Linear Parameters

- The linear parameters have the following property by definition:

$$\lambda^{\hat{P}_n\{P\}\hat{P}_m\{S\}}_{\hat{P}_n\{\nu\}\hat{P}_m\{\mu\}} = \lambda^{\{P\}\{S\}}_{\{\nu\}\{\mu\}}$$

$\boldsymbol{P}_k$ being an arbitrary permutation of k elements

- Thus, all sets of $\{P\}$ and $\{S\}$ which only differ in the ordering are equivalent
- The ordered set of values $\{P\}$, $\{S\}$ is called the **signature** of a particle group
- Each signature corresponds to a unique set of parameters

P and S

- $\{P\}$ are particle-particle types
- $\{P\}$ can be written as a symmetric matrix
  - For a typical electronic calculation we would have:

$$\boldsymbol{P}=\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

- Rows/columns ordered as electron up, electron down
- Distinguishes between parallel- and antiparallel-spin pairs
- Does not distinguish between up-up and down-down

  - For an electron-hole system we could have:

$$\boldsymbol{P}=\begin{pmatrix} \text{١} & 2 & 3 & 3 \\ 2 & 1 & 3 & 3 \\ 3 & 3 & \text{٤} & 5 \\ 3 & 3 & 5 & 4 \end{pmatrix}$$

- Rows/columns ordered as e-up, e-down, h-up, h-down
- Like above within electrons, and within holes
- Distinguishes between electron pairs and hole pairs
- All electron-hole pairs indistinguishable regardless of spin

P and S

- In a term with n=3, m=0 (three-electron), given:

$$P = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

we can have the following combinations of particle types (second arrow means sorting):

  - Types: (1,1,1) → pairs: (1,1,1) → signature (1,1,1)
  - Types: (1,1,2) → pairs: (1,2,2) → signature (1,2,2)
  - Types: (1,2,2) → pairs: (2,2,1) → signature (1,2,2)
  - Types: (2,2,2) → pairs: (1,1,1) → signature (1,1,1)

- Therefore we have two independent sets of linear parameters (**channels**)

P and S

- {S} are particle-nucleus types
- {S} can be written as a rectangular matrix
    - For a water molecule we could have:

$$\boldsymbol{S}=\begin{pmatrix}1 & 1\\ 2 & 3\\ ٢ & 3\end{pmatrix}$$

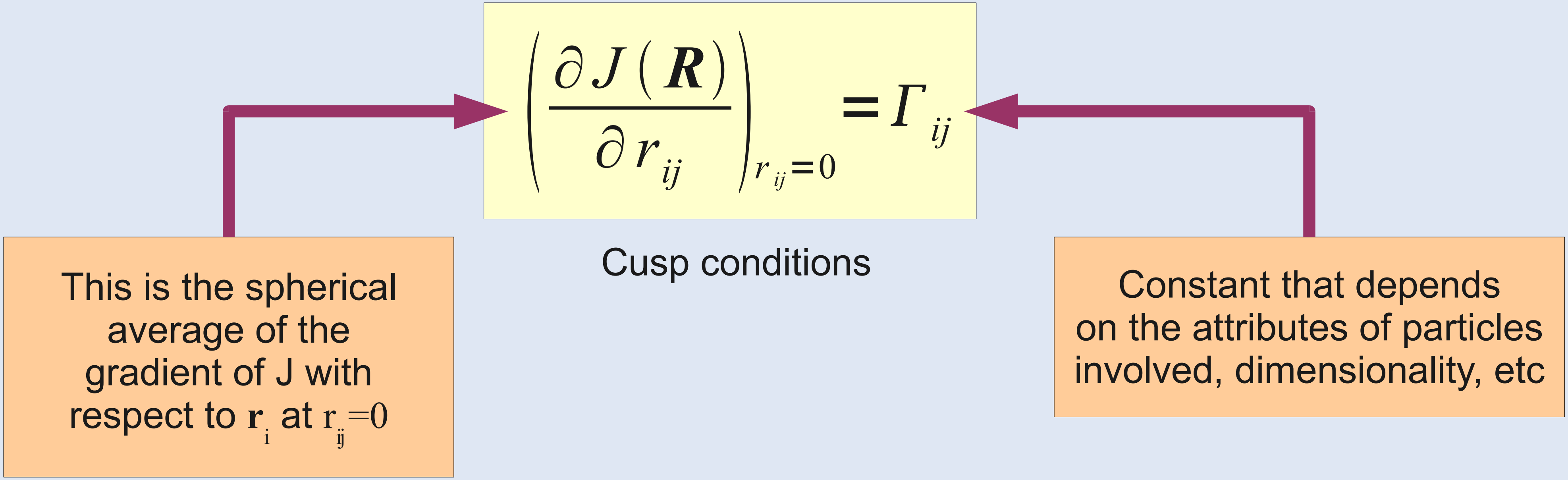
- Rows ordered as O, H, H; columns as e-up, e-down
- Distinguishes between O and H
- Does not distinguish the two H's
- Does not distinguish up- and down-electrons in relation to O
- Distinguishes up- and down-electrons in relation to H

- Full signatures are obtained from both $\{P\}$ and $\{S\}$ (ordering occurs **within** each set, though)
- NB, basis functions depend on a single P_{ij} or S_{iI}: channel splitting is independent from that of linear parameters

Constraints on the Linear Parameters

$$\lambda^{\{P\}\qquad\{S\}}_{\hat{P}_n\{\nu\}\hat{P}_m\{\mu\}} = \lambda^{\{P\}\{S\}}_{\{\nu\}\{\mu\}} \quad \forall\, \hat{P}_n,\, \hat{P}_m :\; \hat{P}_n\{P\}=\{P\},\, \hat{P}_m\{S\}=\{S\}$$

Repeated signature indices imply symmetry with respect to exchange of bottom indices.



$$\left(\frac{\partial J(\boldsymbol{R})}{\partial r_{ij}}\right)_{r_{ij}=0} = \Gamma_{ij}$$

Cusp conditions

This is the spherical average of the gradient of J with respect to $\mathbf{r}_i$ at $r_{ij}=0$

Constant that depends on the attributes of particles involved, dimensionality, etc

Cusp Conditions

- Non-zero cusp only applicable to two-body terms:

$$\left(\frac{\partial J_{2,0}(\boldsymbol{R})}{\partial r_{ij}}\right)_{r_{ij}=0} = \sum_{\nu_{ij}=1}^{p} \lambda_{\nu_{ij}}^{P_{ij}} \left(\frac{\partial \Phi_{\nu_{ij}}^{P_{ij}}(\boldsymbol{r}_{ij})}{\partial r_{ij}}\right)_{r_{ij}=0} = \Gamma_{ij}$$

$$\left(\frac{\partial J_{1,1}(\boldsymbol{R})}{\partial r_{iI}}\right)_{r_{iI}=0} = \sum_{\mu_{iI}=1}^{p} \lambda_{\mu_{iI}}^{S_{iI}} \left(\frac{\partial \Theta_{\mu_{iI}}^{S_{iI}}(\boldsymbol{r}_{iI})}{\partial r_{iI}}\right)_{r_{iI}=0} = \Gamma_{iI}$$

- Of course, only one e-e term should have a non-zero cusp for a given particle-pair type. Same goes for e-n term.

Cusp Conditions

- The particle-particle no-cusp condition is:

$$\underbrace{\sum^{p}}_{\{\nu_{ii_\beta},\,\nu_{ji_\beta}\}_{\beta=3,n}} \underbrace{\sum^{q}}_{\{\mu_{iI_\gamma},\,\mu_{jI_\gamma}\}_{\gamma=1,m}} \left[\sum_{\nu_{ij}}^{p} \lambda_{\{\nu\}\{\mu\}}^{\{P\}\{S\}} \left(\frac{\partial\, \boldsymbol{\Phi}_{\nu_{ij}}^{P_{\sigma(i)\sigma(j)}}\left(\boldsymbol{r}_{ij}\right)}{\partial\, r_{ij}} \right)_{r_{ij}=\cdot} \right] \times$$
$$\times \prod_{\beta=3}^{n} \Phi_{\nu_{ii_\beta}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right) \Phi_{\nu_{ji_\beta}}^{P_{\sigma(j)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right) \prod_{\gamma=1}^{m} \Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right) \Theta_{\mu_{jI_\gamma}}^{S_{\sigma(j)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right) = 0$$

  - The first sum is constrained to index pairs such that $\Phi_{\nu_{ii_\beta}}^{P_{\sigma(i)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right) \Phi_{\nu_{ji_\beta}}^{P_{\sigma(j)\sigma(i_\beta)}}\left(\boldsymbol{r}_{ii_\beta}\right)$ gives the same function
  - The second sum is constrained to index pairs such that $\Theta_{\mu_{iI_\gamma}}^{S_{\sigma(i)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right) \Theta_{\mu_{jI_\gamma}}^{S_{\sigma(j)I_\gamma}}\left(\boldsymbol{r}_{iI_\gamma}\right)$ gives the same function
  - In many common cases, only the square bracket needs to be zero

Cusp Conditions

- The particle-nucleus no-cusp condition is:

$$\sum_{\underbrace{}_{\{\nu_{ii_\alpha},\mu_{i_\alpha J}\}_{\alpha=2,n}}}^{p,q} \sum_{\mu_{iJ}}^{q} \left[\lambda^{\{P\}\{S\}}_{\{\nu\}\{\mu\}} \left(\frac{\partial \Theta^{S_{\sigma(i)J}}_{\mu_{iJ}}\left(\boldsymbol{r}_{iJ}\right)}{\partial r_{iJ}} \right)_{r_{iJ}=0} \right] \times$$

$$\times \prod_{\beta=2}^{n} \Phi^{P_{\sigma(i)\sigma(i_\beta)}}_{\nu_{ii_\beta}}\left(\boldsymbol{r}_{ii_\beta}\right) \Theta^{S_{\sigma(i_\beta)J}}_{\mu_{i_\beta J}}\left(\boldsymbol{r}_{i_\beta J}\right) = 0$$

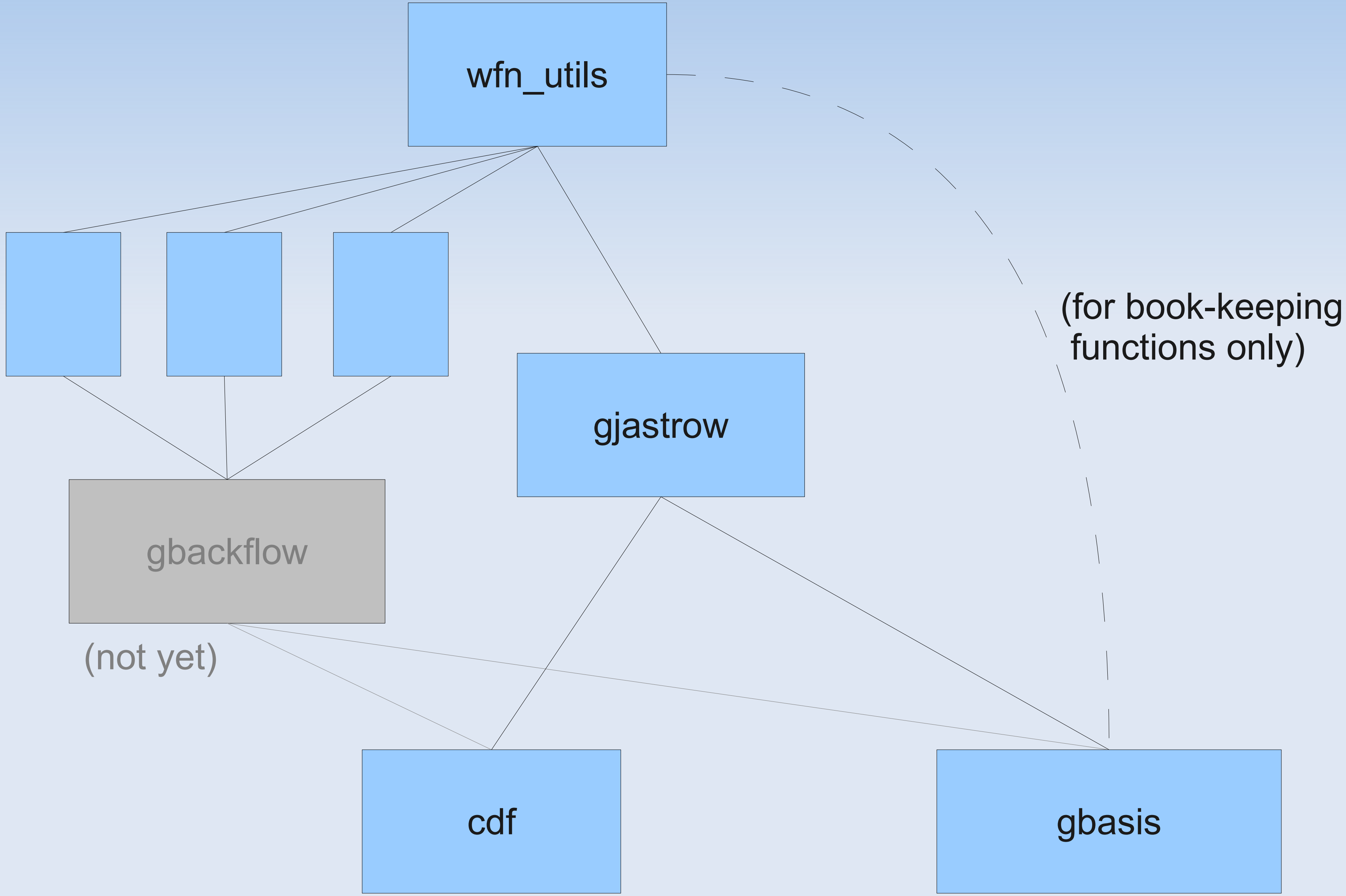
  - The first sum is constrained to index pairs such that $\Phi^{P_{\sigma(i)\sigma(i_\alpha)}}_{\nu_{ii_\alpha}}\left(\boldsymbol{r}_{ii_\alpha}\right) \Theta^{S_{\sigma(i_\alpha)J}}_{\mu_{i_\alpha J}}\left(\boldsymbol{r}_{i_\alpha J}\right)$ gives the same function
  - In many common cases, only the square bracket needs to be zero

Pseudocode for Term Evaluation

```
Get basis and cutoff functions
Loop over ion_vector(1:m)
  Loop over ispin_vector(1:n)
    Get signature and permutation
    Loop over ie_vector(1:n)
      Get cutoff functions (incremental)
      Loop over nu_vector(1:(n*(n-1))/2)
        Get product of e-e basis functions (permuted, incremental)
        Loop over mu_vector(1:n*m)
          Get product of e-n basis functions (permuted, incremental)
          Accumulate contribution
        End loop over mu_vector
      End loop over nu_vector
    End loop over ie_vector
  End loop over ispin_vector
End loop over ion_vector
```

It is convenient to write optimized versions
of the loops, for performance.

Code Structure



Basis-Function Handling

- The gbasis module handles basis and cutoff functions
- Sharing of basis functions among several terms is allowed for functions without parameters in them
- Sharing of basis functions between Jastrow and backflow modules also possible

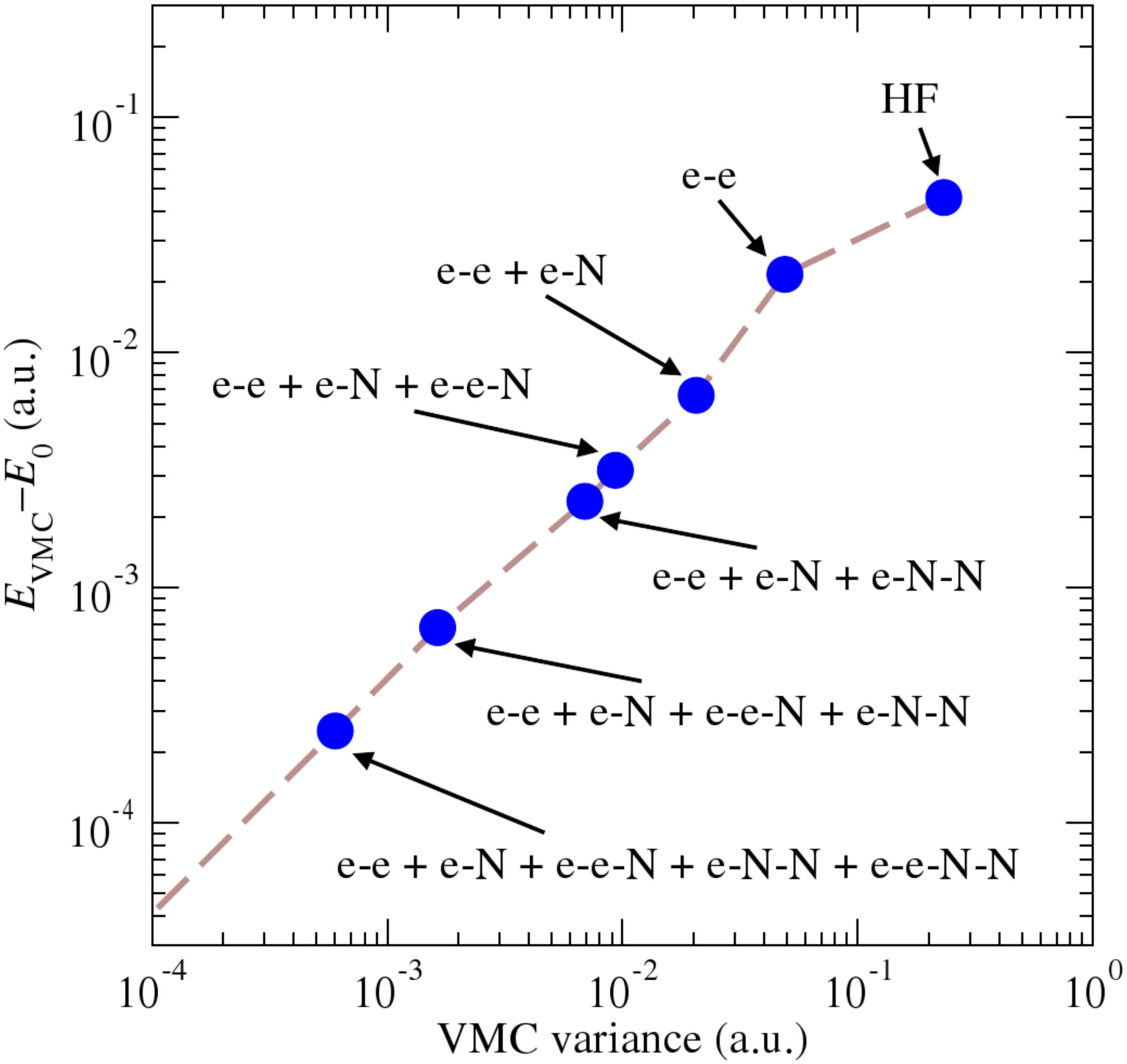
The CASINO Data Format (CDF)

- Format for structured data resembling YAML
- Nodes in the structure may be:
  - scalars (keyword ↔ value, or value only)
  - blocks (keyword ↔ set of children nodes)
- Indentation used to indicate depth
- Alternatively, square brackets and commas can be used for children of a block
- Values are strings, converted to other types on request
- Keywords are case- and whitespace-insensitive

The CASINO Data Format (CDF)

```
JASTROW:
  Title: 2D HEG
  TERM 1:
    Rank: [ 2, 0 ]
    Rules: [ 1-1 = 2-2 ]
    e-e basis: [ Type: polynomial, Order: 9 ]
    e-e cutoff:
      Type: polynomial
      Constants: [ C: 3 ]
      Parameters:
        Channel 1-1: [ L: [ 4.51888387, optimizable ] ]
        Channel 1-2: [ L: [ 4.51888387, optimizable ] ]
    Linear parameters:
      Channel 1-1:
        c_1: [ 5.10744890784656D-003, optimizable ]
        c_2: [ -2.215745836891912E-004, fixed ]
        c_3: [ 2.61049912223271D-003, optimizable ]
        c_4: [-5.71501823949904D-003, optimizable ]
        c_5: [ 6.26629894651862D-003, optimizable ]
        c_6: [-3.69279305571898D-003, optimizable ]
        c_7: [ 1.21802169144595D-003, optimizable ]
        c_8: [-2.11189983328774D-004, optimizable ]
        c_9: [ 1.50406603295831D-005, optimizable ]
```

H_2 molecule


$E_{VMC} - E_0$ (a.u.)
10^{-1}
10^{-2}
10^{-3}
10^{-4}
HF
e-e
e-e + e-N
e-e + e-N + e-e-N
e-e + e-N + e-N-N
e-e + e-N + e-e-N + e-N-N
e-e + e-N + e-e-N + e-N-N + e-e-N-N
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
VMC variance (a.u.)

Current State of the Project

- CDF → done
- Channel splitting → done
- Basis set evaluation → done (basic ones)
- Basis set sharing → done
- Term evaluation → done
- Initial tests → done
- Speed → done
- Constraints → done
- Tests on interesting systems → in progress