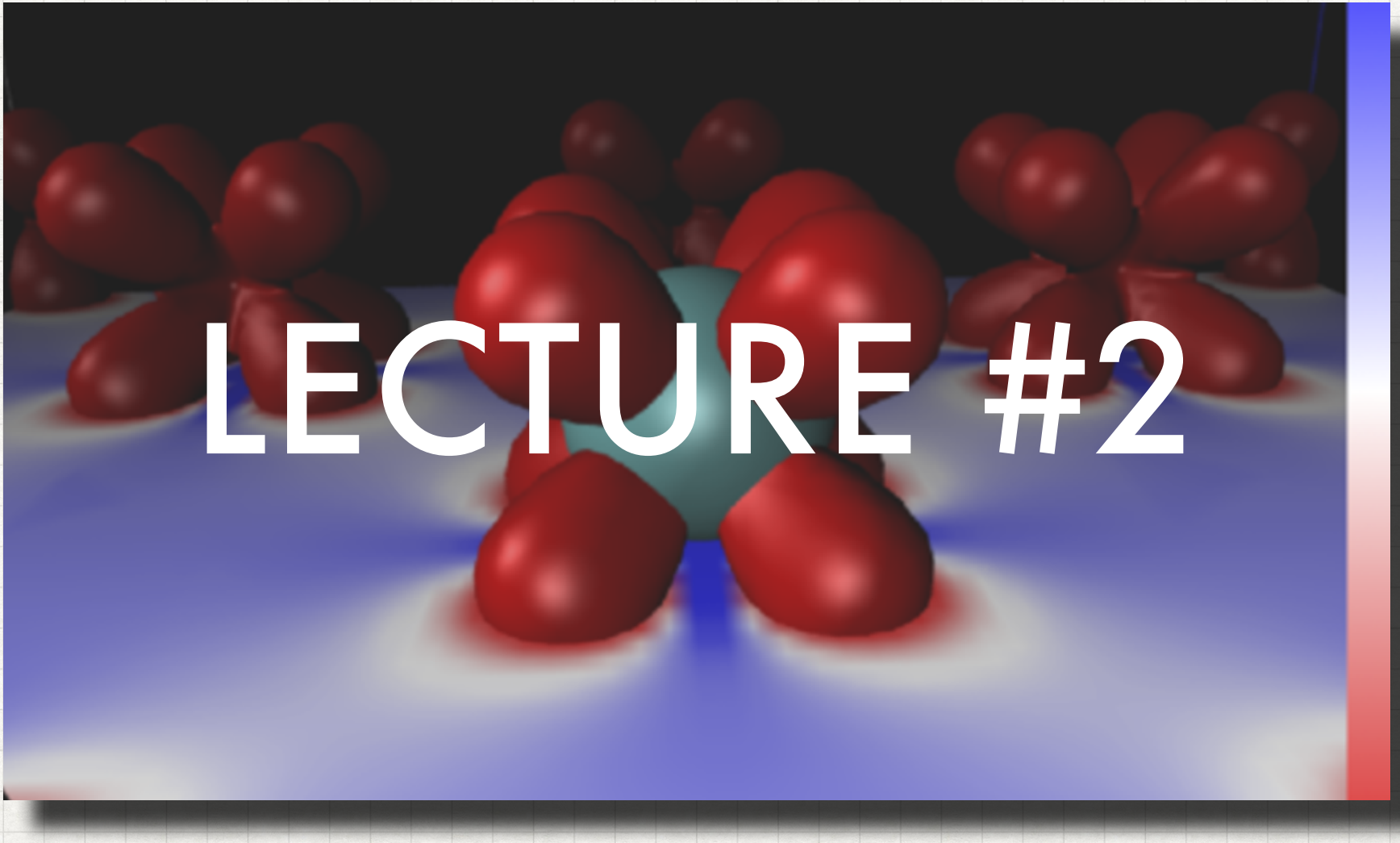
COUPLED CLUSTER THEORY

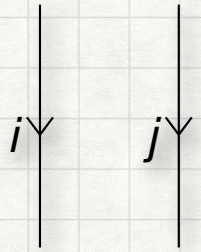
PROF. T. DANIEL CRAWFORD, VIRGINIA TECH



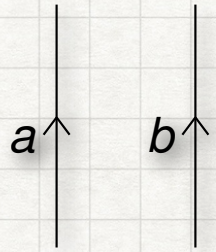
Coester and Kümmel adopted the name "exp-S method", after the form of the wave function (2) and the symbol S for the cluster operation. Cizek and Paldus introduced the term "Coupled-Pair Many Electron Theory" (CPMET), and extended CPMET, from their emphasis on $T \simeq T_2$, explicit equations for which were first obtained by Cizek [3a]. We prefer the denotation Coupled-Clutter Method (CCM) since it suggests most vividly the central features of admissibility of general-sized clusters and coupling of them in a nonlinear fashion. It is worth pointing out that Eq. (28) represents not a unitary but rather a similarity transformation. This is unlike the situation in work like that of Westhaus [8], as it voids the termination of the commutator series but preserves the Hermiticity of the transformed Hamiltonian.

H.J. Monkhorst, "Calculation of properties with the coupled-cluster method",
Int. J. Quantum Chem. Symp., 11, 421-432 (1977).

DIAGRAMS: THE BASICS

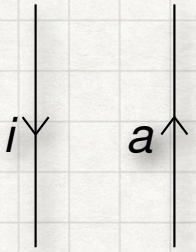


Holes
(Occupied Orbitals)



Particles
(Virtual Orbitals)

The Fermi Vacuum
(Empty Space)

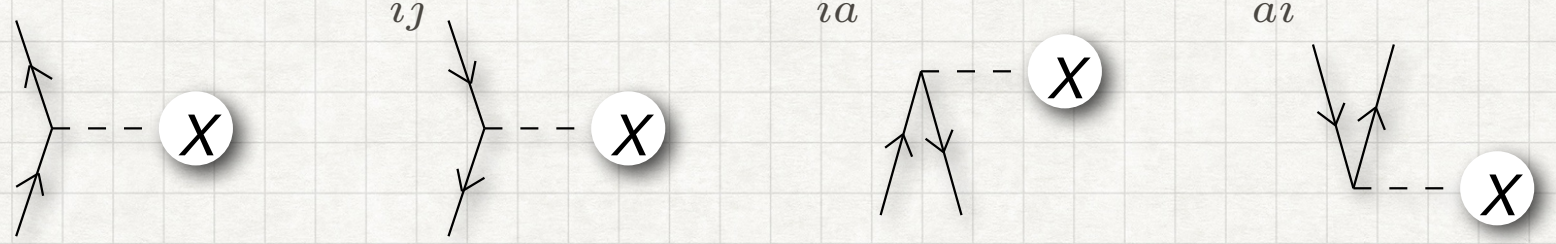


A Singly Excited
Determinant, $|\Phi_i^a\rangle$

DIAGRAMS: THE BASICS

- Diagrammatic representation of the Fock operator

$$\hat{F}_N = \sum_{ab} f_{ab}\{a_a^\dagger a_b\} + \sum_{ij} f_{ij}\{a_i^\dagger a_j\} + \sum_{ia} f_{ia}\{a_i^\dagger a_a\} + \sum_{ai} f_{ai}\{a_a^\dagger a_i\}$$



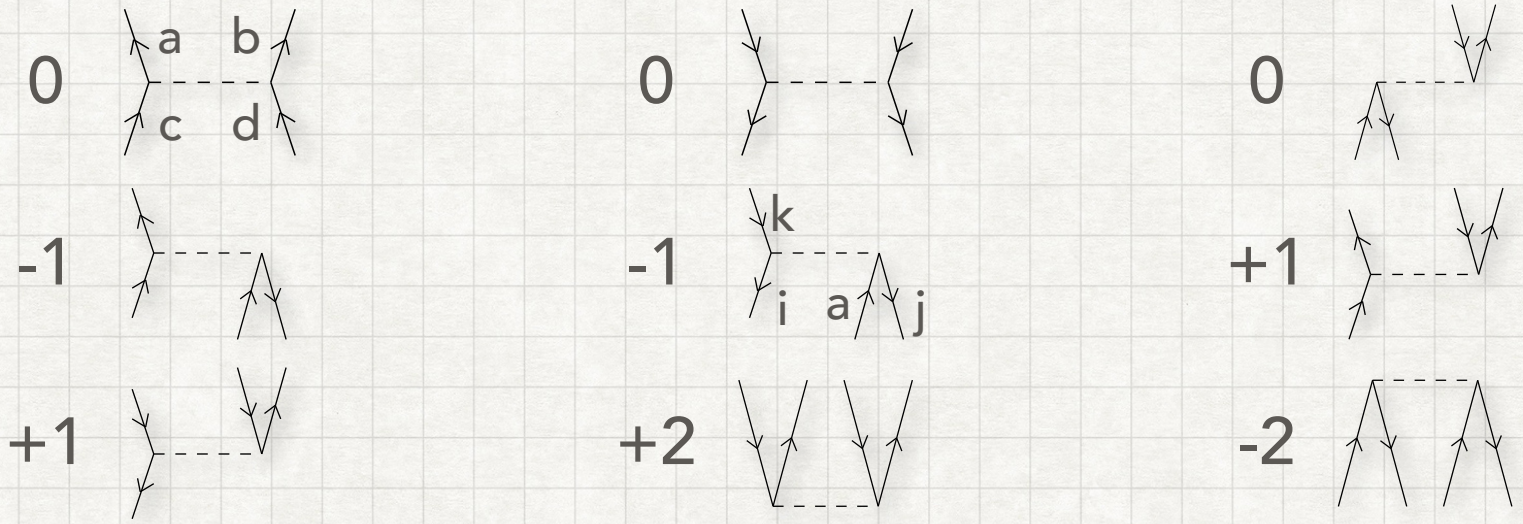
Excitation Level: 0 0 -1 +1

- The dotted horizontal line is called the "interaction line", and the upward- and downward-directed lines emanate from "vertices" and represent the annihilation/creation operators. Each vertex can connect to two directed lines, at most.
- The outward-directed line from a vertex corresponds to the left-hand orbital index in the integral, and the inward-directed line to the right-hand index.
- Directed lines above the interaction line are "quasiparticle creation operators" ($a_a^\dagger$ or a_i), and those below are "quasiparticle annihilation operators" (a_a or $a_i^\dagger$). The excitation level of a diagram is the difference in the number of creation and annihilation operators divided by 2.

DIAGRAMS: THE BASICS

- Diagrammatic representation of the two-electron Hamiltonian component:

$$\hat{V}_N = \frac{1}{4}\sum_{abcd}\langle ab||cd\rangle\{a_a^\dagger a_b^\dagger a_d a_c\} + \frac{1}{4}\sum_{ijkl}\langle ij||kl\rangle\{a_i^\dagger a_j^\dagger a_l a_k\} + \sum_{iabj}\langle ia||bj\rangle\{a_i^\dagger a_a^\dagger a_j a_b\}$$
$$+ \frac{1}{2}\sum_{aibc}\langle ai||bc\rangle\{a_a^\dagger a_i^\dagger a_c a_b\} + \frac{1}{2}\sum_{ijka}\langle ij||ka\rangle\{a_i^\dagger a_j^\dagger a_a a_k\} + \frac{1}{2}\sum_{abci}\langle ab||ci\rangle\{a_a^\dagger a_b^\dagger a_i a_c\}$$
$$+ \frac{1}{2}\sum_{iajk}\langle ia||jk\rangle\{a_i^\dagger a_a^\dagger a_k a_j\} + \frac{1}{4}\sum_{abij}\langle ab||ij\rangle\{a_a^\dagger a_b^\dagger a_j a_i\} + \frac{1}{4}\sum_{ijab}\langle ij||ab\rangle\{a_i^\dagger a_j^\dagger a_b a_a\}$$



- For two-electron operators/integrals, the ordering of indices is:

⟨left-out right-out||left-in right-in⟩

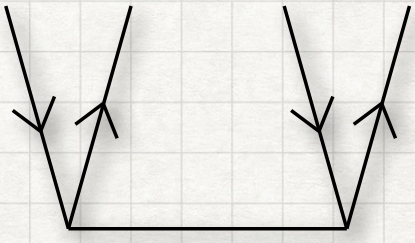
DIAGRAMS: THE BASICS

- Diagrammatic representation of cluster operators:

$$\hat{T}_1 = \sum_{ia} t_i^a \{a_a^\dagger a_i\}$$

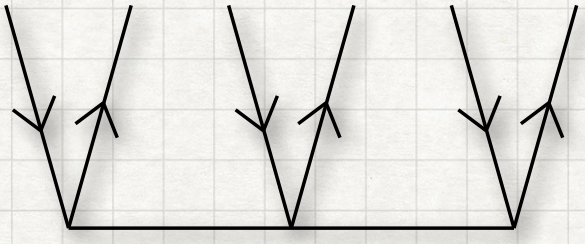
+1

$$\hat{T}_2 = \frac{1}{4} \sum_{ijab} t_{ij}^{ab} \{a_a^\dagger a_b^\dagger a_j a_i\}$$



+2

$$\hat{T}_3 = \frac{1}{36} \sum_{ijkabc} t_{ijk}^{abc} \{a_a^\dagger a_b^\dagger a_c^\dagger a_k a_j a_i\}$$



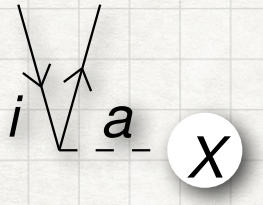
+3

- We choose a solid (rather than dotted) interaction line to distinguish the $\hat{T}_n$ from $\hat{V}_N$.
- All lines are quasiparticle creation lines in this case because these are excitation operators.

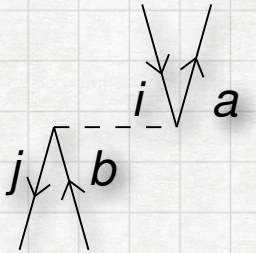
DIAGRAMS: MATRIX ELEMENTS

- We usually interpret our diagrams as matrix elements of second-quantized operators (or products of such operators) between Slater determinants.
- Right-to-left in a matrix element corresponds to bottom-to-top in a given diagram.
- Examples:

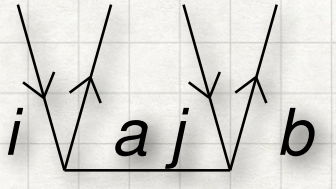
$$\langle \Phi_i^a | \hat{F}_N | \Phi_0 \rangle =$$



$$\langle \Phi_i^a | \hat{V}_N | \Phi_j^b \rangle =$$



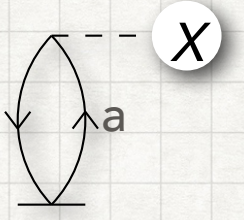
$$\langle \Phi_{ij}^{ab} | \hat{T}_2 | \Phi_0 \rangle =$$



DIAGRAMS: THE CC ENERGY

- We are now ready to construct diagrams representing the components of the coupled cluster energy equation. We'll start with the simplest term that is linear in $\hat{T}_1$:

$$\langle \Phi_0 | \left(\hat{F}_N \hat{T}_1 \right)_c | \Phi_0 \rangle =$$



$$= \sum_{ia} \sum_{ia} f_i^a t_i^a$$

- To interpret this algebraically:
  1. Label all directed lines with appropriate indices: *i*, *j*, *k*, …; *a*, *b*, *c*, …
  2. Include summations for all "internal" lines (those that begin and end at interaction lines). "External" lines start at an interaction line and extend above of below the diagram.
  3. Interpret the lines connected to the Fock operator and $\hat{T}_1$ vertices as:

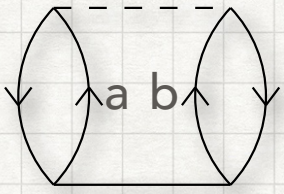
$$f^{\text{left-out}}_{\text{right-in}} \qquad t^{\text{out}}_{\text{in}}$$

  4. The sign of the diagram is $(-1)^{h+l}$, where *h* is the number of hole lines and *l* is the number of "loops" – a route along a series of directed lines that either returns to its beginning or starts at an external line and ends at another.

DIAGRAMS: THE CC ENERGY

- Next consider the terms linear in $\hat{T}_2$:

$$\langle \Phi_0 | \left(\hat{H}_N \hat{T}_2 \right)_c | \Phi_0 \rangle =$$



$$= \frac{1}{4} \sum_{ijab} \langle ij||ab \rangle t_{ij}^{ab}$$

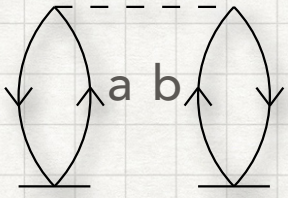
- To interpret this algebraically:
  1. Label all directed lines with appropriate indices: *i*, *j*, *k*, …; *a*, *b*, *c*, …
  2. Include summations for all "internal" lines.
  3. Interpret the lines connected to the $\hat{V}_N$ operator vertex as:

     $\langle$left-out right-out$||$left-in right-in$\rangle$
  4. Interpret the lines connected to the $\hat{T}_2$ operator from left to right:

     $t^{\text{left right}}_{\text{left right}}$
  5. The sign of the diagram is $(-1)^{h+l}$.
  6. Paris of "equivalent" lines begin at the same interaction line and end at the same interaction line. For each such pair, multiply the expression by 1/2.

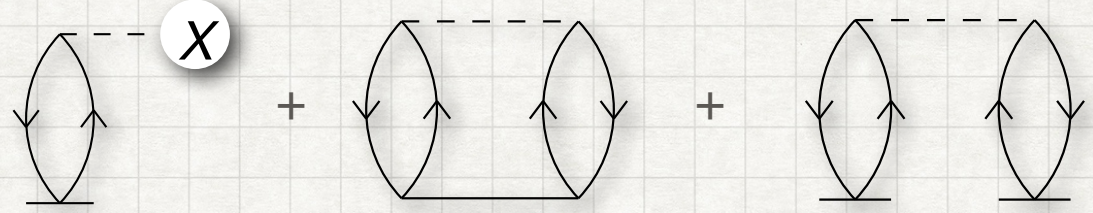
DIAGRAMS: THE CC ENERGY

- Next consider the terms quadratic in $\hat{T}_1$:

$$\langle \Phi_0 | \left(\frac{1}{2} \hat{H}_N \hat{T}_1^2 \right)_c | \Phi_0 \rangle =$$ 

$$= \frac{1}{2} \sum_{ijab} \langle ij||ab \rangle t_i^a t_j^b$$

- This diagram may be interpreted using the previous rules, as appropriate, but now we have one additional rule:
  - A pair of "equivalent vertices" is given when two identical operators connect to the Hamiltonian in exactly the same way. For *n* equivalent vertices, multiply the expression by *1/n!*.
- Putting all three diagrams together gives us our final expression:

$$E_{CCSD} =$$ 

$$= \sum_{ia} f_{ia} t_i^a + \frac{1}{4} \sum_{ijab} \langle ij||ab \rangle t_{ij}^{ab} + \frac{1}{2} \sum_{ijab} \langle ij||ab \rangle t_i^a t_j^b$$

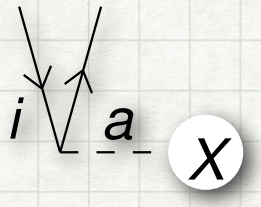
DIAGRAMS: THE CC AMPLITUDE EQUATIONS

- Next we consider the amplitude equations:

$$0 = \langle \Phi_i^a | \left(\hat{H}_N e^{\hat{T}} \right)_c | \Phi_0 \rangle \quad \text{and} \quad 0 = \langle \Phi_{ij}^{ab} | \left(\hat{H}_N e^{\hat{T}} \right)_c | \Phi_0 \rangle$$

- The leading term in the $\hat{T}_1$ equation is:

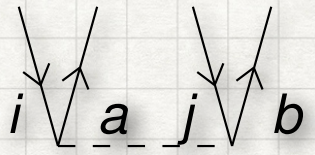
$$\langle \Phi_i^a | \hat{F}_N | \Phi_0 \rangle =$$



$$= f_{ai}$$

- The leading term in the $\hat{T}_2$ equation is:

$$\langle \Phi_{ij}^{ab} | \hat{V}_N | \Phi_0 \rangle =$$

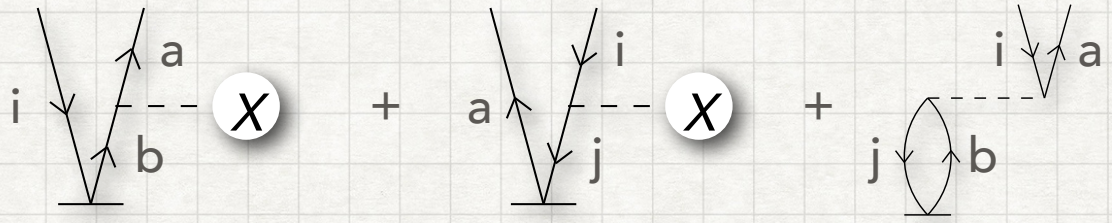


$$= \langle ab || ij \rangle$$

DIAGRAMS: THE CC AMPLITUDE EQUATIONS

- The linear $\hat{T}_1$ contributions to the singles equation are:

$$\langle \Phi_i^a | \left(\hat{F}_N \hat{T}_1 \right)_c | \Phi_0 \rangle + \langle \Phi_i^a | \left(\hat{V}_N \hat{T}_1 \right)_c | \Phi_0 \rangle =$$

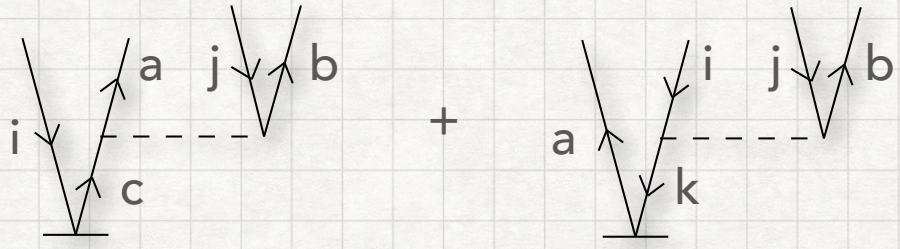


$$= \sum_b f_{ab} t_i^b - \sum_j f_{ji} t_j^a + \sum_{jb} \langle ja || bi \rangle t_j^b$$

DIAGRAMS: THE CC AMPLITUDE EQUATIONS

- The linear $\hat{T}_1$ contribution of $\hat{V}_N$ to the doubles equation is:

$$\langle \Phi_{ij}^{ab} | \left(\hat{V}_N \hat{T}_1 \right)_c | \Phi_0 \rangle =$$



- For these diagrams, we need a new rule:

Each pair of unique (*i.e.*, connecting to different interaction lines), external hole or particle lines introduces a permutation operator, $P(pq)$, to ensure antisymmetry of the final expression.

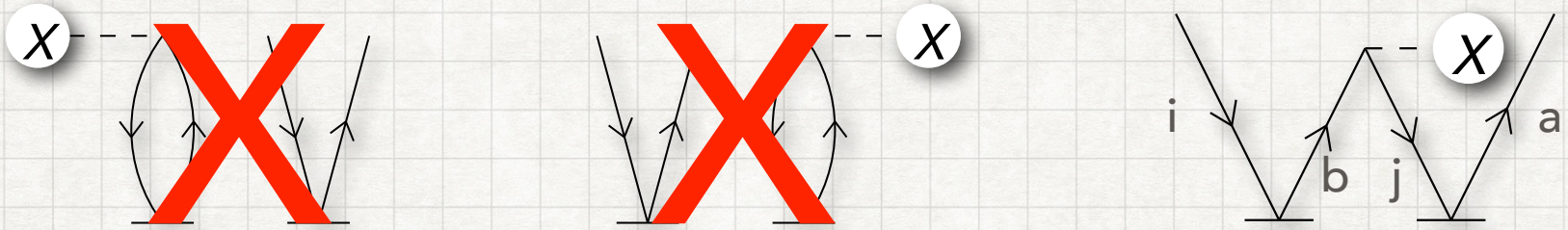
- This rule, in conjunction with our previous rules, gives:

$$\langle \Phi_{ij}^{ab} | \left(\hat{V}_N \hat{T}_1 \right)_c | \Phi_0 \rangle = P(ij) \sum_c \langle ab || cj \rangle t_i^c - P(ab) \sum_k \langle kb || ij \rangle t_k^a$$

DIAGRAMS: THE CC AMPLITUDE EQUATIONS

- The quadratic $\hat{T}_1$ contribution of $\hat{F}_N$ to the singles equation is:

$$\frac{1}{2}\langle\Phi_i^a|\left(\hat{F}_N\hat{T}_1^2\right)_c|\Phi_0\rangle =$$



- Only those diagrams in which the Hamiltonian fragment has at least one connection/contraction with every $\hat{T}$ operator can contribute.

$$\frac{1}{2}\langle\Phi_i^a|\left(\hat{F}_N\hat{T}_1^2\right)_c|\Phi_0\rangle = -\sum_{jb} f_{jb}t_i^b t_j^a$$

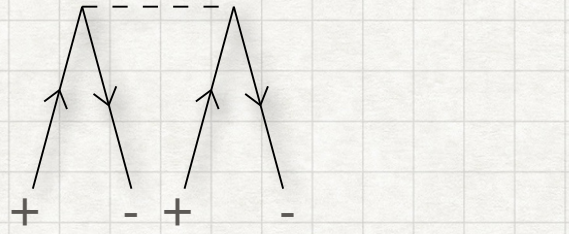
- Note that the factor of 1/2 in the expression on the left is automatically accounted for in the diagram.

DIAGRAMS: THE CC AMPLITUDE EQUATIONS

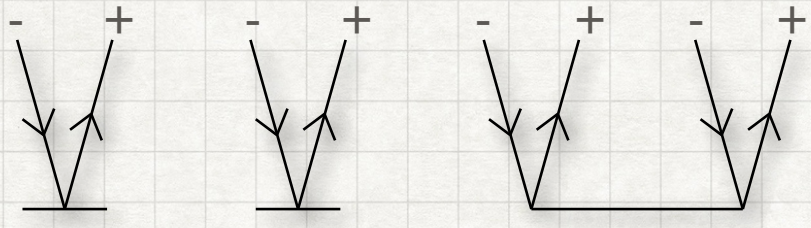
- How can we be sure that we include only the ***unique*** diagrams? Consider the following contribution to the $\hat{T}_2$ equations:

$$\frac{1}{2}\langle\Phi_{ij}^{ab}|\left(\hat{F}_N\hat{T}_1^2\hat{T}_2\right)_c|\Phi_0\rangle$$

- We know that the product of $\hat{T}_1^2$ and $\hat{T}_2$ yields a +4 excitation, and the matrix element requires a total +2 excitation (the doubly excited determinant on the left), so only the -2 diagram from the Hamiltonian will work:



-2



+4

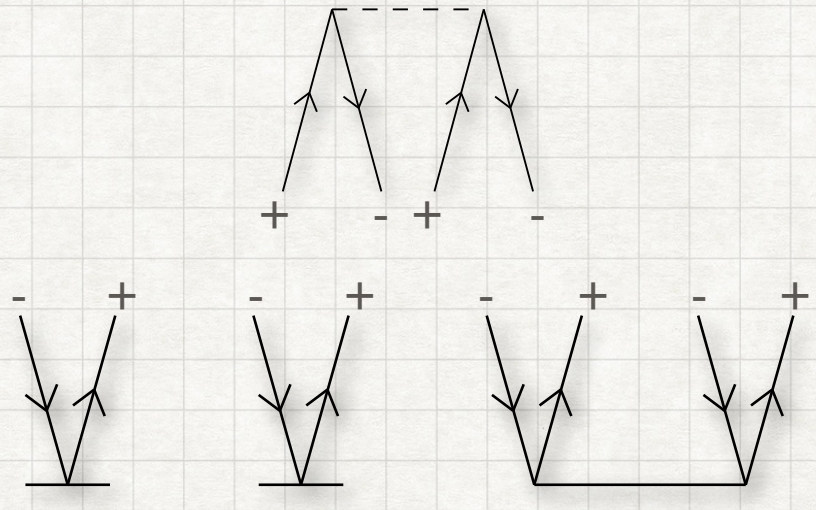
- The Kucharski-Bartlett "sign sequence": Assign "+" to each particle line and "-" to each hole line from all operators and select all unique combinations that take at least one connection from each $\hat{T}$ operator.

DIAGRAMS: THE CC AMPLITUDE EQUATIONS

$$\frac{1}{2}\langle\Phi_{ij}^{ab}|\left(\hat{F}_N\hat{T}_1^2\hat{T}_2\right)_c|\Phi_0\rangle$$

- Since there can be only four connections (+-+-), we choose up to four +/- signs from each of the three cluster operators:

$\hat{T}_1$	$\hat{T}_1$	$\hat{T}_2$
+	+	– –
–	–	+ +
+	–	+ –
+ –	+	–
+ –	–	+

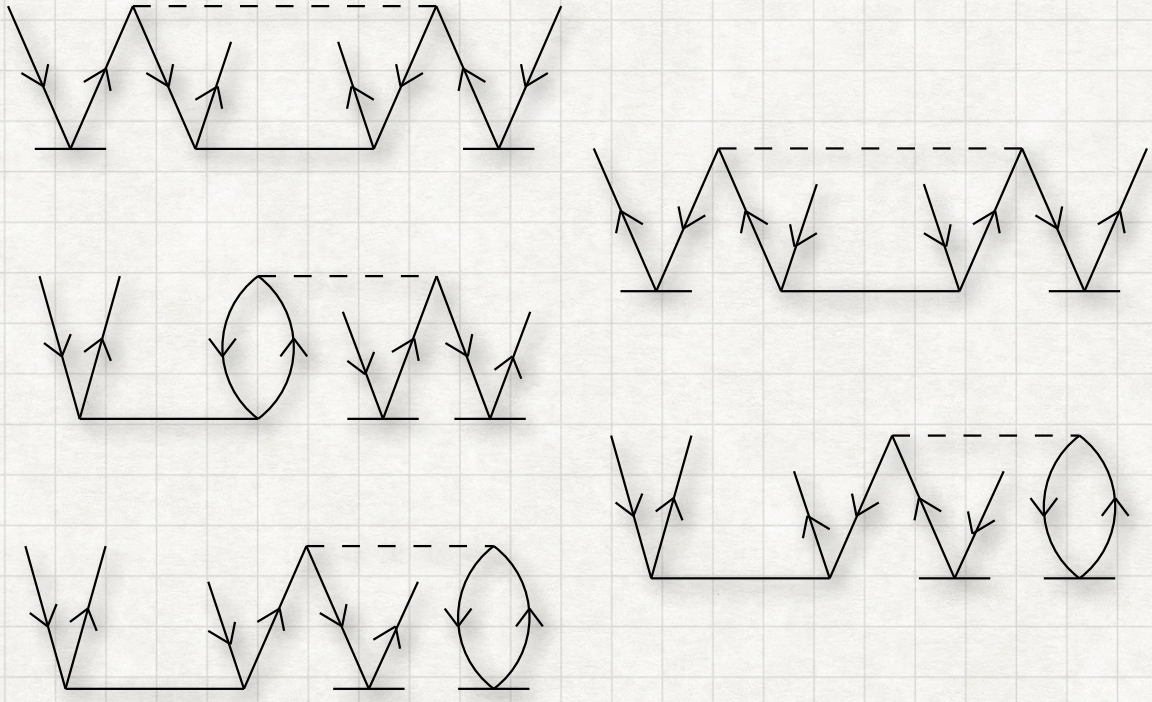


- There are only five unique combinations, leading to five unique diagrams.

DIAGRAMS: THE CC AMPLITUDE EQUATIONS

$$\frac{1}{2}\langle\Phi_{ij}^{ab}|\left(\hat{F}_N\hat{T}_1^2\hat{T}_2\right)_c|\Phi_0\rangle$$

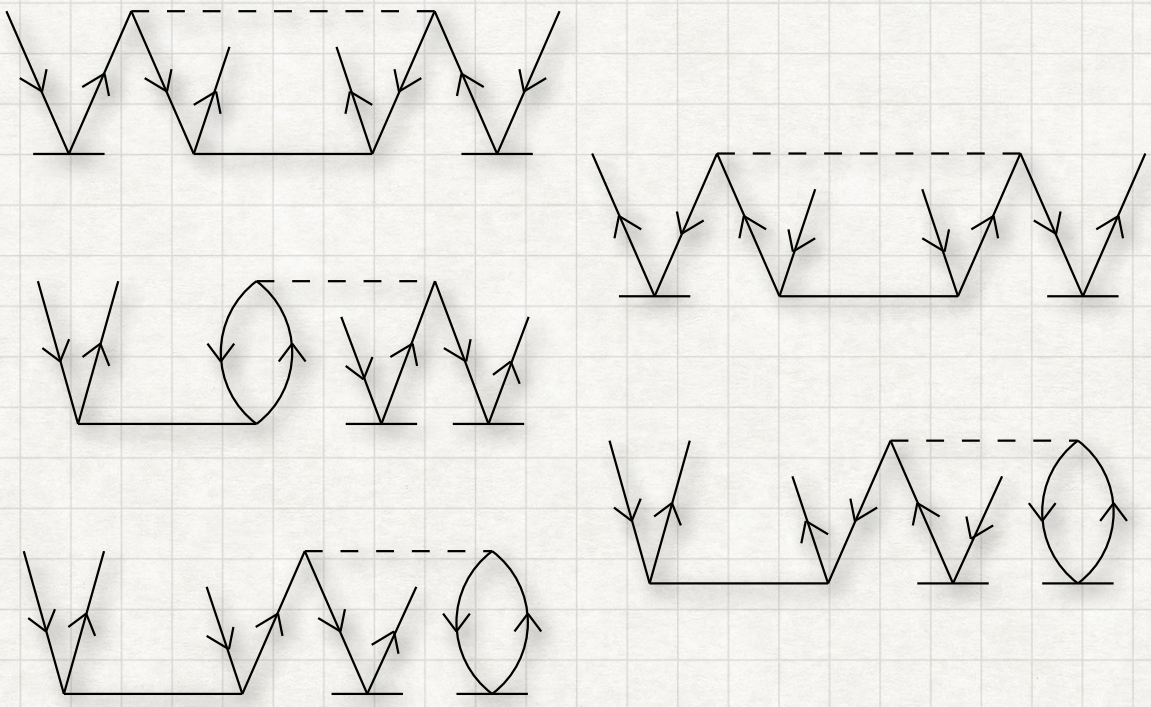
$\hat{T}_1$	$\hat{T}_1$	$\hat{T}_2$
+	+	– –
–	–	+ +
+	–	+ –
+ –	+	–
+ –	–	+



- There are only five unique combinations, leading to five unique diagrams.

DIAGRAMS: THE CC AMPLITUDE EQUATIONS

$$\frac{1}{2}\langle\Phi_{ij}^{ab}|\left(\hat{F}_N\hat{T}_1^2\hat{T}_2\right)_c|\Phi_0\rangle$$



- Applying our rules, we have the complete algebraic interpretation:

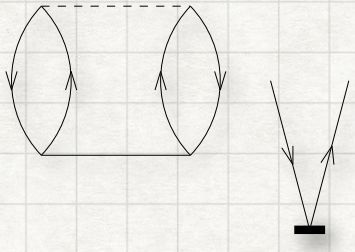
$$\frac{1}{2}\langle\Phi_{ij}^{ab}|\left(\hat{V}_N\hat{T}_1^2\hat{T}_2\right)_c|\Phi_0\rangle = +\frac{1}{4}P(ij)\sum_{klcd}\langle kl||cd\rangle t_{kl}^{ab}t_i^c t_j^d + \frac{1}{4}P(ab)\sum_{klcd}\langle kl||cd\rangle t_{ij}^{cd}t_k^a t_l^b$$
$$- P(ij)P(ab)\sum_{klcd}\langle kl||cd\rangle t_{ik}^{ac}t_j^d t_l^b - P(ij)\sum_{klcd}\langle kl||cd\rangle t_{ik}^{ab}t_j^c t_l^d - P(ab)\sum_{klcd}\langle kl||cd\rangle t_{ij}^{ac}t_k^b t_l^d$$

SIZE EXTENSIVITY

- This term was developed by Bartlett in 1978[a] and refers to the correct physical (linear) scaling of the energy with the number of electrons.

Configuration Interaction	Coupled Cluster
$E = \langle \Phi_0 \vert \hat{H} \left(1 + \hat{C}\right) \vert \Phi_0 \rangle$	$E = \langle \Phi_0 \vert e^{-\hat{T}} \hat{H} e^{\hat{T}} \vert \Phi_0 \rangle$
$E c_{ij\ldots}^{ab\ldots} = \langle \Phi_{ij\ldots}^{ab\ldots} \vert \hat{H} \left(1 + \hat{C}\right) \vert \Phi_0 \rangle$	$0 = \langle \Phi_{ij\ldots}^{ab\ldots} \vert e^{-\hat{T}} \hat{H} e^{\hat{T}} \vert \Phi_0 \rangle$

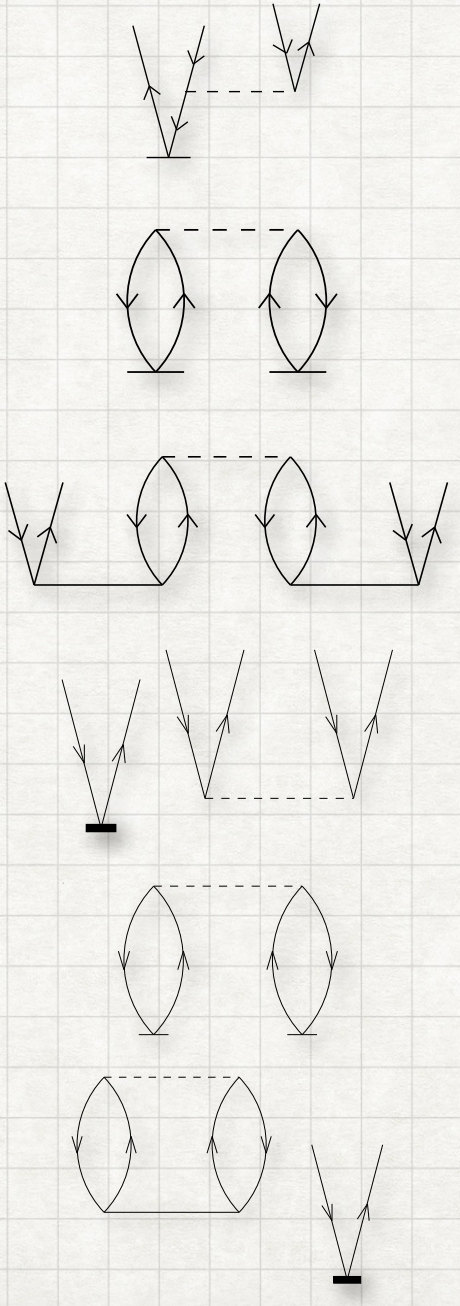
- The (unavoidable) appearance of the energy in the CI amplitude equations leads to “unlinked" diagrams such as:



- The scaling of the "closed" component of the diagram, which represents the CI energy, scales independently from the open component, leading to non-linear scaling of the CI energy with the size of the system.
- Such terms cannot appear in the coupled cluster equations, and thus the CC energy scales correctly (linearly) with the number of electrons.

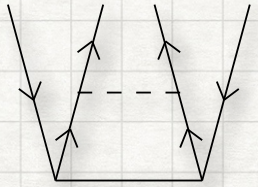
COMMON DIAGRAM TERMINOLOGY

- ***Open*** – contains external lines
- ***Closed*** – no external lines
- ***Connected*** – all parts of the diagram are connected to one another via directed lines
- ***Disconnected*** – some components of the diagram are not connected via directed lines
- ***Linked*** – contains closed components that are connected
- ***Unlinked*** – contains closed components that are not connected



COMPUTATIONAL COST

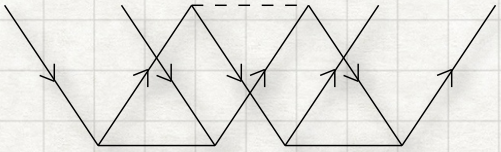
- Now that we have algebraic and diagrammatic representations of the many terms of the CC equations, we may examine their computational cost in detail. The most expensive term in the CCSD equations is the particle-particle ladder:



$$= \frac{1}{2}\sum_{ef}\langle ab||ef\rangle t_{ij}^{ef} \qquad \text{Cost:} \quad \mathcal{O}(n_o^2 n_v^4)$$

This is typically simplified to $\mathcal{O}(N^6)$

- However, if you look closely at the $\hat{T}_2$ amplitude equations we saw earlier, you'll find terms that appear to be even more expensive, *e.g.*:



$$= \frac{1}{4}\sum_{mnef} t_{ij}^{ef} t_{mn}^{ab} \langle mn||ef\rangle \qquad \mathcal{O}(n_o^4 n_v^4)$$

- Such terms, which involve products of $\hat{T}$ operators, may be factorized into less expensive terms:

$$\frac{1}{4}\sum_{mnef} t_{ij}^{ef} t_{mn}^{ab} \langle mn||ef\rangle = \frac{1}{4}\sum_{mn} t_{mn}^{ab} \left(\sum_{ef} t_{ij}^{ef} \langle mn||ef\rangle\right) \qquad 2 \times \mathcal{O}(n_o^4 n_v^2)$$