Simple formalism for efficient derivatives and multi-determinant expansions in Quantum Monte Carlo

Bastien Mussard, **Julien Toulouse**, **Roland Assaraf**

Laboratoire de Chimie Théorique
Institut du Calcul et de la Simulation
Sorbonne Universités, Université Pierre et Marie Curie

bastien.mussard@upmc.fr
www.lct.jussieu.fr/pagesperso/mussard/

Needs

(Note : the derivatives involved are often "logarithmic derivatives" : $\frac{\partial_\mu A}{A} = \partial_\mu \ln A$)

First-order derivatives of the wavefunction : $\frac{\partial_\mu \Psi}{\Psi}$

- wrt. electronic coordinates (drift)
- wrt. nuclear coordinates (internuclei forces)
- wrt. parameters of the wavefunction (optimizations)

One-body operators : $\frac{\hat{O}\Psi}{\Psi}$

- local quantities, ...

Derivatives of local quantities : $\partial_\mu \frac{\hat{O}\Psi}{\Psi}$

- ...

Update of determinants :

- determinant are modified when electrons move
- excited determinants differ from a reference determinant by a few columns

First-order derivatives $\left(\frac{\partial_\mu \det(A)}{\det(A)}\right)$ and One-body operators $\left(\frac{\hat{O}\det(A)}{\det(A)}\right)$

We will show that they are the same.
(A can be any matrix, we will consider the Slater matrix.)

First-order derivatives of a determinant

$$\boxed{\frac{\partial_\mu \det(A)}{\det(A)} = \mathrm{tr}\left(A^{-1}\partial_\mu A\right)} \tag{1}$$

Element of proof:

- A chain rule on the logarithmic derivative gives :

$$\partial_\mu \ln\det(A) = \sum_I \frac{\partial \ln\det(A)}{\partial A_{ij}}\partial_\mu A_{ij}$$

and an expansion of the determinant in minor will gives :

$$\frac{\partial \ln\det(A)}{\partial A} = A^{-1}$$

Alternatively :

- $\ln\det(A) = \mathrm{tr}(\ln A)$, and : $\partial_\mu \ln A = \frac{\partial_\mu A}{A}$

One-body operations on a determinant

$$\boxed{\frac{\hat{O}\det(A)}{\det(A)} = \mathrm{tr}\left(A^{-1}B\right)} \qquad B_{ij} = (O\phi_j)(\mathbf{r}_i) \tag{2}$$

Element of proof:

- Consider the operator : $\hat{O} = O(\mathbf{r}_1) + O(\mathbf{r}_2) + \cdots + O(\mathbf{r}_N)$.
  It's action on $\det(A)$ is :

$$\begin{aligned}\hat{O}\det(A) &= \sum (-1)^p \Big(O(\mathbf{r}_1) + O(\mathbf{r}_2) + .. + O(\mathbf{r}_N)\Big)\Big(\phi_1(\mathbf{r}_{p(1)})\phi_2(\mathbf{r}_{p(2)})..\phi_N(\mathbf{r}_{p(N)})\Big) \\ &= |(O\phi_1)\phi_2\ldots\phi_N| + |\phi_1(O\phi_2)\ldots\phi_N| + \cdots + |\phi_1\phi_2\ldots(O\phi_N)|\end{aligned}$$

- Consider $\det(A_\lambda) = \det(A+\lambda B) = |(\phi_1+\lambda O\phi_1)(\phi_2+\lambda O\phi_2)\ldots(\phi_N+\lambda O\phi_N)|$
  It's derivative wrt λ, taken at $\lambda = 0$, is (multilinearity of "det") :

$$\partial_\lambda \det(A_\lambda)|_{\lambda=0} = |(O\phi_1)\phi_2\ldots\phi_N| + |\phi_1(O\phi_2)\ldots\phi_N| + |\phi_1\phi_2\ldots(O\phi_N)|$$

- Eq. (1) : $\dfrac{\hat{O}\det(A)}{\det(A)} = \left.\dfrac{\partial_\lambda \det(A_\lambda)}{\det(A_\lambda)}\right|_{\lambda=0} = \mathrm{tr}\left(A_\lambda^{-1}\partial_\lambda A_\lambda\right)|_{\lambda=0} = \mathrm{tr}\left(A^{-1}B\right)$

Note : we use consistent notations to separate
actual derivatives ($\partial_\mu \Phi$, with "μ"s) from trick derivatives ($\partial_\lambda d_\lambda|_{\lambda=0} = \hat{O}\,d$, with "$\lambda$"s)

One-body operators as first-order derivatives

The equations (1) and (2), that are repeated here :

$$\frac{\partial_\mu \det(A)}{\det(A)} = \mathrm{tr}\left(A^{-1}\partial_\mu A\right) \qquad \frac{\hat{O}\det(A)}{\det(A)} = \mathrm{tr}\left(A^{-1}B\right)$$

show that, formally, **any one-body operator can be written using the same compact expression as first-order derivatives**.

Take the example of the kinetic operator $\hat{T} = -\frac{1}{2}\Delta$: the Laplacian is the sum of **second order derivatives** with respect to the electron coordinates, but can be written as a first-order derivative when applied to a Slater determinant (with $B_{ij} = -\frac{1}{2}\Delta\phi_j(\mathbf{r}_i)$).

This is the basis of an important simplification in the computation of all local quantity and their derivatives, **as everything that will be derived is valid for both derivatives and one-body operators** : with $B_{ij} = (O\phi_j)(\mathbf{r}_i)$.

We will talk about "operations on determinants" for both first-order derivatives and one-body operations.

In the following, the notations involving B assume that an operator $\hat{O}$ is applied to the elements of the matrix with the same dressing as B : B_I is $\hat{O}A_I$, $\tilde{B}$ is $\hat{O}\tilde{A}$, etc. . .

Notations

Notations that will simplify the derivations of updates of determinants and inverses.

Notations (1/2)

- The wavefunction of an N-electron system is $\Psi = J\,\Phi = J\,\sum_I c_I \det(A_I)$

  where J is a Jastrow, and $\det(\boldsymbol{A_I})$ are determinants of matrices A_I.
  The matrices A_I differ from a reference matrix A of determinant $\det(\boldsymbol{A})$ by $\boldsymbol{k_I}$ columns.

- The columns are picked from the $N \times N_{\text{tot}}$ extended Slater matrix $\tilde{\boldsymbol{A}}$ that contains all orbitals :

$$\tilde{A} = \begin{pmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \dots & \phi_N(\mathbf{r}_1) & \phi_{N+1}(\mathbf{r}_1) & \dots & \phi_{N_{\text{tot}}}(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & \dots & \phi_N(\mathbf{r}_2) & \phi_{N+1}(\mathbf{r}_2) & \dots & \phi_{N_{\text{tot}}}(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & \dots & \phi_N(\mathbf{r}_N) & \phi_{N+1}(\mathbf{r}_N) & \dots & \phi_{N_{\text{tot}}}(\mathbf{r}_N) \end{pmatrix}$$

- The $N_{\text{tot}} \times N$ matrix $\boldsymbol{R_I}$ selects (when applied on the right) the N orbitals from $\tilde{A}$ that compose the matrix A_I, so that $A_I = \tilde{A} R_I$, in particular : $A = \tilde{A} R$.

Notations (2/2)

- Since all that is needed is the knowledge of the k_I columns that change between A_I and the reference A, we define $\boldsymbol{P_I}$, the $N \times k_I$ matrix that selects those columns (when applied on the right) : $A_I P_I$ is the $N \times k_I$ matrix made of the columns that changed with respect to A (see example on next slide).

  With $i_1, i_2, \ldots, i_k$ the ordered indexes of the columns that are modified :

$$(P_I)_{i_1,1} = (P_I)_{i_2,2} = \cdots = (P_I)_{i_k,k} = 1, \qquad (P_I)_{ij} = 0 \quad \text{otherwise}$$

  Note that $P_I P_I^T$ is a $N \times N$ diagonal matrix with ones on positions that correspond to a column that changes :

$$(P_I P_I^T)_{ij} = \sum_k (P_I)_{ik}(P_I)_{jk} = \begin{cases} \delta_{ij} & \text{if } i \in \{i_1, i_2, \ldots, i_k\} \\ 0 & \text{otherwise} \end{cases}$$

  Consequently, the identity $A_I - A = (A_I - A) P_I P_I^T$ holds.

- Since $A_I P_I$ are columns picked from $\tilde{A}$, we can also define $\boldsymbol{Q_I}$, the $N_{\text{tot}} \times k_I$ matrix so that $A_I P_I = \tilde{A} Q_I$ (see example on next slide).

  A relation between R_I and R then emerges :

$$A_I = A + (A_I - A) P_I P_I^T = \tilde{A}(R + (Q_I - R P_I) P_I^T) = \tilde{A} R_I$$

Example (1/2)

- Consider a **4-electron** system described by a total of **7 orbitals**.

$$\tilde{A} = \begin{pmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 & \phi_6 & \phi_7 \\ \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 & \phi_6 & \phi_7 \\ \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 & \phi_6 & \phi_7 \\ \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 & \phi_6 & \phi_7 \end{pmatrix}$$

- The **reference matrix** can for example be :

$$A = \tilde{A}R = \begin{pmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 \\ \phi_1 & \phi_2 & \phi_3 & \phi_4 \\ \phi_1 & \phi_2 & \phi_3 & \phi_4 \\ \phi_1 & \phi_2 & \phi_3 & \phi_4 \end{pmatrix} \quad \text{where} \quad R = \begin{array}{c} \text{number of the} \\ 1^{\text{st}} \quad 2^{\text{nd}} \quad 3^{\text{rd}} \quad 4^{\text{th}} \\ \text{orbital} \\ \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

- The **double-excitation** $\phi_2 \rightarrow \boldsymbol{\phi_5}$ and $\phi_4 \rightarrow \boldsymbol{\phi_7}$ is :

$$A_I = \tilde{A}R_I = \begin{pmatrix} \phi_1 & \boldsymbol{\phi_5} & \phi_3 & \boldsymbol{\phi_7} \\ \phi_1 & \boldsymbol{\phi_5} & \phi_3 & \boldsymbol{\phi_7} \\ \phi_1 & \boldsymbol{\phi_5} & \phi_3 & \boldsymbol{\phi_7} \\ \phi_1 & \boldsymbol{\phi_5} & \phi_3 & \boldsymbol{\phi_7} \end{pmatrix} \quad \text{where } R_I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{pmatrix}$$

Example (2/2)

- Consider a **4-electron** system described by a total of **7 orbitals**.

$$\tilde{A} = \begin{pmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 & \phi_6 & \phi_7 \\ \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 & \phi_6 & \phi_7 \\ \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 & \phi_6 & \phi_7 \\ \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 & \phi_6 & \phi_7 \end{pmatrix} \quad ; \quad A_I = \begin{pmatrix} \phi_1 & \boldsymbol{\phi_5} & \phi_3 & \boldsymbol{\phi_7} \\ \phi_1 & \boldsymbol{\phi_5} & \phi_3 & \boldsymbol{\phi_7} \\ \phi_1 & \boldsymbol{\phi_5} & \phi_3 & \boldsymbol{\phi_7} \\ \phi_1 & \boldsymbol{\phi_5} & \phi_3 & \boldsymbol{\phi_7} \end{pmatrix}$$

- The **selector matrices** read :

$$P_I = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

There are 2 excitations (two columns), and they are on the second and fourth columns of A_I.

$$A_I P_I = \tilde{A} Q_I = \begin{pmatrix} \boldsymbol{\phi_5} & \boldsymbol{\phi_7} \\ \boldsymbol{\phi_5} & \boldsymbol{\phi_7} \\ \boldsymbol{\phi_5} & \boldsymbol{\phi_7} \\ \boldsymbol{\phi_5} & \boldsymbol{\phi_7} \end{pmatrix}$$

$$Q_I = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \mathbf{1} & 0 \\ 0 & 0 \\ 0 & \mathbf{1} \end{pmatrix}$$

There are 2 excitations (two columns), and they concern the **fifth** and **seventh** orbitals (columns of $\tilde{A}$).

Operations on Φ $\left(\frac{\hat{O}\Phi}{\Phi}\right)$

$\Phi = \sum_I c_I \det(A_I)$ is composed of determinants

- operations on Φ, involves a matrix Γ
- efficient computation of Γ (via update of determinants and inverses)

(At the end are presented comparisons with Petruzielo and Umrigar notes.)

Operations on the CI expansion

$$\boxed{\frac{\hat{O}\Phi}{\Phi} = \Phi^{-1}\,\mathrm{tr}\left(\Gamma\,\tilde{B}\right)} \qquad \Gamma = \sum_I c_I\,\det\left(A_I\right) R_I\,A_I^{-1} \tag{3}$$

Element of proof:

$$\begin{aligned}
\hat{O}\Phi &= \sum c_I\;\hat{O}\det\left(A_I\right) \\
&= \sum c_I\,\det\left(A_I\right)\,\mathrm{tr}\left(A_I^{-1}B_I\right) && \left(\text{Eq. (1)} : \hat{O}\det\left(A\right) = \det\left(A\right)\,\mathrm{tr}\left(A^{-1}B\right)\right) \\
&= \sum c_I\,\det\left(A_I\right)\,\mathrm{tr}\left(A_I^{-1}\tilde{B}R_I\right) && \left(\text{remember} : A_I = \tilde{A}R_I\right)
\end{aligned}$$

Computing Γ (step 1 : excited determinant $\det(A_I)$ updated from $\det(A)$)

$$\boxed{\det(A_I) = \det(A)\,\det(U_I)} \qquad U_I = P_I^T A^{-1} \tilde{A} Q_I \tag{4}$$

Element of proof: (this is the determinant lemma of Sherman-Morrison-Woodbury)

$$\begin{aligned}
\det(A_I) &= \det\left(A + (A_I - A) P_I P_I^T\right) \\
&= \det(A)\det\left(1 + P_I^T A^{-1}(A_I - A) P_I\right) \\
&= \det(A)\det\left(P_I^T A^{-1} A_I P_I\right) \qquad \text{(and remember : } A_I P_I = \tilde{A} Q_I\text{)}
\end{aligned}$$

This is the first occurence of $A^{-1}\tilde{A}$, a very important matrix in the following.

For a k-th order excitation I, one only need to compute the determinant of U_I, a $k \times k$ submatrix of $A^{-1}\tilde{A}$.

Note that the block composed of the first N columns of $A^{-1}\tilde{A}$ is the identity matrix, and is never used. In practice, one needs only to compute the $N \times N_{\text{vir}}$ submatrix, and actually only the N_s elements $(A^{-1}\tilde{A})_{ij}$ corresponding to monoexcitations $i \to j$ that are used in the CI expansion. (Note that in [Clark *et al.* 2011], the matrix T is such a subset of $A^{-1}\tilde{A}$.)

In our example, we merely have : $U_I = \begin{pmatrix} (A^{-1}\tilde{A})_{2,5} & (A^{-1}\tilde{A})_{2,7} \\ (A^{-1}\tilde{A})_{4,5} & (A^{-1}\tilde{A})_{4,7} \end{pmatrix}$

Computing Γ (step 2 : excited inverse A_I^{-1} updated from A^{-1})

$$R_I A_I{}^{-1} = RA^{-1} + (1 - RA^{-1}\tilde{A})\, Q_I\, U_I^{-1}\, P_I^T\, A^{-1}$$

Element of proof:

- The Sherman-Morrison-Woodbury formula for inverses is, with our notations :

$$\begin{aligned} A_I^{-1} &= A^{-1} + A^{-1}(AP_I - A_I P_I)\, U_I^{-1}\, P_I^T A^{-1} \\ &= A^{-1} + (P_I - A^{-1}\tilde{A} Q_I)\, U_I^{-1}\, P_I^T A^{-1} \end{aligned} \qquad (\text{remember : } A_I P_I = \tilde{A} Q_I)$$

- Multiplying from the left by P_I^T, this gives :

$$\begin{aligned} P_I^T A_I^{-1} &= P_I^T A^{-1} + (P_I^T P_I - U_I)\, U_I^{-1}\, P_I^T A^{-1} \qquad (U_I = P_I^T A^{-1} \tilde{A} Q_I) \\ &= P_I^T P_I\, U_I^{-1}\, P_I^T A^{-1} = U_I^{-1}\, P_I^T A^{-1} \quad \leftarrow \text{(b)} \end{aligned}$$

- Remembering that $R_I = (R + (Q_I - RP_I)P_I^T)$, we have :

$$\begin{aligned} R_I A_I^{-1} &= \left(R + (Q_I - RP_I)P_I^T\right)\left(A^{-1} + (P_I - A^{-1}\tilde{A} Q_I)\, U_I^{-1}\, P_I^T A^{-1}\right) \\ &= RA^{-1} + RP_I\, U_I^{-1}\, P_I^T A^{-1} - RA^{-1}\tilde{A} Q_I\, U_I^{-1}\, P_I^T A^{-1} \\ &\quad + Q_I P_I^T A^{-1} + Q_I P_I^T P_I\, U_I^{-1}\, P_I^T A^{-1} - Q_I P_I^T A^{-1} \tilde{A} Q_I\, U_I^{-1}\, P_I^T A^{-1} \\ &\quad - RP_I P_I^T A_I^{-1} \end{aligned}$$

The blue terms cancel thanks to (b), the red terms cancel since $U_I = P_I^T A^{-1} \tilde{A} Q_I$.

Relation with the notes of Petruzielo and Umrigar

In those notes we have : $A_I = A + UV$, i.e. $U = (A_I - A)P_I$ and $V = P_I^T$.

Hence, with the formula in those notes

$$\det(A_I) = \det(A)\det\left(I_k + VA^{-1}U\right) = \det(A)\det(D)$$

we correctly recover $D = I_k + P_I^T A^{-1}(A_I - A)P_I = P_I^T A^{-1} A_I P_I = U_I$

The relation $[D]_{ij} = [A^{-1}A_I]_{f(i),f(j)}$, which is a little bit convoluted to demonstrate in those notes, appear here **directly from the actions of P_I^T and P_I** : indeed, with our example of P_I, the action on any matrix is :

$$P_I^T M P_I = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{2,2} & a_{2,4} \\ a_{4,2} & a_{4,4} \end{pmatrix}$$

i.e. it **selects the pertinent informations**, which was the purpose of f in the notes.

Conversely, the action of P_I and P_I^T is to redistribute the information at the correct places in the big matrices :

$$P_I M P_I^T = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_{2,2} & a_{2,4} \\ a_{4,2} & a_{4,4} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a_{2,2} & 0 & a_{2,4} \\ 0 & 0 & 0 & 0 \\ 0 & a_{4,2} & 0 & a_{4,4} \end{pmatrix}$$

Computing Γ

$$\boxed{\Gamma = \Phi R A^{-1} + \det(A)\ (1 - R A^{-1}\tilde{A})\Upsilon A^{-1}} \qquad \Upsilon = \sum_I c_I \det(U_I)\, Q_I\, U_I^{-1}\, P_I^T \quad (5)$$

Element of proof:

$$\begin{aligned}\Gamma &= \sum c_I \det(A_I)\, R_I\, A_I^{-1} \\ &= \sum c_I \,(\det(A)\ \det(U_I))(RA^{-1} + (1 - RA^{-1}\tilde{A})\, Q_I\, U_I^{-1}\, P_I^T\, A^{-1})\end{aligned}$$

On the computation of Υ and Φ :

- When computing Υ, it is better to **gather excitations of the same order** :

$$\Upsilon = \sum_k \Upsilon_k = \sum_k \sum_{I/k_I=k} c_I \det(U_I)\ Q_I\, U_I^{-1}\, P_I^T$$

- To compute Φ, **one does not need to redo the sum over those excitations** :

$$\begin{aligned}\Phi = \sum_I c_I \det(U_I) &= \sum_I c_I \det(U_I)\ \mathrm{tr}\left(U_I^{-1} U_I\right)/k_I \\ &= \sum_k \sum_{I/k_I=k} c_I \det(U_I)\ \mathrm{tr}\left(U_I^{-1} P_I^T A^{-1} \tilde{A} Q_I\right)/k \\ &= \mathrm{tr}\left(A^{-1}\tilde{A} \sum_k \Upsilon_k/k\right)\end{aligned}$$

Summary

Assuming Γ is available, **any first-order derivatives** (drift, internuclei forces, optimization quantities, ...) as well as **any one-body local quantity** (local energy, ...) can all be computed according to :

$$\hat{O}\Phi = \mathrm{tr}\left(\Gamma B\right) \quad \text{where} \quad B_{ij} = (O\phi_j)(\mathbf{r}_i)$$

The cost of computing Γ is

Second-order derivatives

Second-order derivatives

$$\partial_\mu \left(\frac{\hat{O}\Phi}{\Phi} \right) = -\Phi^{-2}\mathrm{tr}\left(\Gamma \partial_\mu \tilde{A}\right)\mathrm{tr}\left(\Gamma \tilde{B}\right) + \Phi^{-1}\mathrm{tr}\left(\partial_\mu \Gamma \tilde{B}\right) + \Phi^{-1}\mathrm{tr}\left(\Gamma \partial_\mu \tilde{B}\right) \tag{6}$$

Element of proof: consider the following development, and everything known :

$$\partial_\mu \left(\Phi^{-1}\mathrm{tr}\left(\Gamma \tilde{B}\right)\right) = \left[-\Phi^{-2}\partial_\mu \Phi\right]\mathrm{tr}\left(\Gamma \tilde{B}\right) + \Phi^{-1}\mathrm{tr}\left(\left[\partial_\mu \Gamma\right]\tilde{B}\right) + \Phi^{-1}\mathrm{tr}\left(\Gamma\left[\partial_\mu \tilde{B}\right]\right)$$

The only new quantity to be computed is $\partial_\mu \Gamma$, which contains $\partial_\mu \Upsilon$. Their expression are too involved to show here, but their cost is the same as Γ and Υ.