

Maxime Chupin

chupin@ceremade.dauphine.fr

www.ceremade.dauphine.fr/~chupin

CNRS, CEREMADE, Université Paris-Dauphine

10 mai 2022

Convergence de la méthode DIIS

appliquée au calcul de configurations électroniques

*en collaboration avec Mi-Song Dupuy, Guillaume Legendre et Éric Séré
(ERC MDFT de M. Lewin)*

Le problème

1 Le problème

2 DIIS et accélération d'Anderson-Pulay

3 Analyse de convergence de variantes de DIIS

4 Résultats numériques

5 Et ensuite...

- **Calcul numérique** de l'état fondamental d'un système moléculaire à N électrons

$$\inf \left\{ \langle \psi, H_v \psi \rangle \mid \psi \in \mathcal{H}, \int_{\mathbb{R}^{3N}} \sum_{\sigma_1, \dots, \sigma_N} |\psi|^2(x_1, \dots, x_N, \sigma_1, \dots, \sigma_N) dx_1 \dots dx_N = 1 \right\}$$

avec

$$H_v = -\frac{1}{2} \sum_{i=1}^N \Delta_{x_i} + \sum_{1 \leq i < j \leq N} \frac{1}{\|x_i - x_j\|} + \sum_{i=1}^N v(x_i)$$

et

$$\mathcal{H} = \bigwedge_{i=1}^N H^1(\mathbb{R}^3 \times \{|+\rangle, |-\rangle\}, \mathbb{C}).$$

- **Problème variationnel** qui ne peut être traité « directement » : **coût de calcul trop élevé** pour N plus grand que quelques unités

- Problème discrétisé (cas **RHF**) :

$$\inf \{E^{\text{RHF}}(D), D \in \mathcal{P}_N\}$$

avec $D = CC^*$ **matrice de densité**, $C \in \mathbb{R}^{d \times N}$ matrice des coefficients orbitaux, $\mathcal{P}_N = \{D \in \mathcal{M}_{d,d}(\mathbb{C}), D^* = D, DSD = D, \text{tr}(SD) = N\}$, et S matrice de Gram.

- Équations d'Euler–Lagrange :

$$\begin{cases} F(D)C = SCA \\ C^*SC = I_N \\ D = CC^* \end{cases}$$

$F(D)$ **matrice de Fock**, Λ multiplicateurs de Lagrange.

- ▶ Reformulation d'une **condition nécessaire** pour que D_* soit solution

$$[F(D_*), D_*] = f(D_*) = 0,$$

où $[A, B] = ABS - SBA$

- ▶ **Caractérisation de D_*** (principe Aufbau) :

$$D_* = \arg \inf \{ \text{tr}(F(D_*)D'), D' \in \mathcal{P}_N \} = g(D_*)$$

DIIS et accélération d'Anderson-Pulay

1 Le problème

2 DIIS et accélération d'Anderson-Pulay

3 Analyse de convergence de variantes de DIIS

4 Résultats numériques

5 Et ensuite...

Direct Inversion in the Iterative Space (Pulay, 1980)

- ▶ Méthode de point fixe

$$\forall k \in \mathbb{N}, x^{(k+1)} = g(x^{(k)}),$$

pour $f(x_*) = 0$

- ▶ DIIS : accélération de la méthode de point fixe. Soit $m \in \mathbb{N}$ (fixe),

$$\forall k \in \mathbb{N}, x^{(k+1)} = \sum_{i=0}^{m_k} c_i^{(k)} g(x^{(k-m_k+i)}),$$

avec $m_k = \min\{m, k\}$ et tel que

$$\left\| \sum_{i=0}^{m_k} c_i^{(k)} f(x^{(k-m_k+i)}) \right\|_2 \text{ est minimal sous la contrainte } \sum_{i=0}^{m_k} c_i^{(k)} = 1.$$

- ▶ Accélération d'Anderson :

$$\forall k \in \mathbb{N}, x^{(k+1)} = g(x^{(k)}) + \sum_{i=1}^{m_k} \theta_i^{(k)} \left(g(x^{(k-m_k-1+i)}) - g(x^{(k)}) \right),$$

avec

$$\theta^{(k)} = \arg \min_{(\theta_1, \dots, \theta_{m_k}) \in \mathbb{R}^{m_k}} \left\| f(x^{(k)}) + \sum_{i=1}^{m_k} \theta_i \left(f(x^{(k-m_k-1+i)}) - f(x^{(k)}) \right) \right\|_2$$

- ▶ Équivalence avec DIIS :

$$c_i^{(k)} = \theta_{i+1}^{(k)}, \quad i = 0, \dots, m_k - 1, \quad c_{m_k}^{(k)} = 1 - \sum_{j=1}^{m_k} \theta_j^{(k)}$$

Analyse de convergence de variantes de DIIS

1 Le problème

2 DIIS et accélération d'Anderson-Pulay

3 Analyse de convergence de variantes de DIIS

4 Résultats numériques

5 Et ensuite...

Analyse de convergence de variantes de DIIS

- ▶ Rappel : $x^{(k+1)} = \sum_{i=0}^{m_k} c_i^{(k)} g(x^{(k-m_k+i)})$, CDIIS : $x^{(k+1)} = g(\sum_{i=0}^{m_k} c_i^{(k)} x^{(k-m_k+i)})$
- ▶ g à valeurs dans une sous-variété régulière de \mathbb{R}^n , $g(x) \neq x + f(x)$, $f : \mathbb{R}^n \mapsto \mathbb{R}^p$
- ▶ f et g régulières (\mathcal{C}^2), f non dégénérée en x_* et g contractante au voisinage de x_*

Analyse de convergence de variantes de DIIS

- ▶ Rappel : $x^{(k+1)} = \sum_{i=0}^{m_k} c_i^{(k)} g(x^{(k-m_k+i)})$, CDIIS : $x^{(k+1)} = g(\sum_{i=0}^{m_k} c_i^{(k)} x^{(k-m_k+i)})$
- ▶ g à valeurs dans une sous-variété régulière de \mathbb{R}^n , $g(x) \neq x + f(x)$, $f : \mathbb{R}^n \mapsto \mathbb{R}^p$
- ▶ f et g régulières (\mathcal{C}^2), f non dégénérée en x_* et g contractante au voisinage de x_*

Mécanismes limitant le nombre d'itérés conservés

- 1 **redémarrage (R-CDIIS)** : étant donné $0 < \tau < 1$, $m_{k+1} = 0$ si

$$\tau \|f(x^{(k+1)}) - f(x^{(k-m_k)})\|_2 > \|(\text{Id} - \Pi_k)(f(x^{(k+1)}) - f(x^{(k-m_k)}))\|_2$$

où Π_k est la projection orthogonale sur $\text{Vect}\{f(x^{(k-m_k+i)}) - f(x^{(k-m_k)}), i = 1, \dots, m_k\}$,

- 2 **fenêtre glissante adaptative (AD-CDIIS)** : étant donné $0 < \delta < 1$, m_{k+1} est le plus grand entier inférieur ou égal à $m_k + 1$ tel que

$$\forall i \in \{k+1 - m_{k+1}, \dots, k\}, \delta \|f(x^{(i)})\|_2 < \|f(x^{(k+1)})\|_2$$

Théorème (R-CDIIS)

Soient $0 < \tau < 1$, et $K < \mu < 1$. Si $x^{(0)}$ est suffisamment proche de x_* , alors, tant que $\|f(x^{(k)})\|_2 > 0$, on a

$$m_k \leq \min(k, p)$$

$$\|c^{(k)}\|_\infty < \infty,$$

et

$$\|f(x^{(k+1)})\|_2 \leq \mu^{m_k+1} \|f(x^{(k-m_k)})\|_2.$$

La méthode est donc linéairement r -convergente.

Théorème (AD-CDIIS)

Soient $0 < \delta < K$, et $K < \mu < 1$. Si $x^{(0)}$ est suffisamment proche de x_* , alors, tant que $\|f(x^{(k)})\|_2 > 0$, on a

$$m_k \leq \min(k, p)$$

$$\|c^{(k)}\|_\infty < \infty,$$

et

$$\|f(x^{(k+1)})\|_2 \leq \mu \|f(x^{(k)})\|_2.$$

La méthode est donc linéairement q -convergente.

Corollaire (r-convergence superlinéaire)

- 1 *Supposons que le paramètre de redémarrage à l'itération k soit donné par $\tau^{(k)} = h_1(\|f(x^{(k-m_k)})\|_2)$. Alors,*

$$\forall k \in \mathbb{N}, k - m_k \geq 1, \quad \|f(x^{(k)})\|_2 \leq C_1 \mu^{m_k} \|f(x^{(k-m_k-m_{k-m_k})})\|_2^{1+\epsilon}.$$

- 2 *Supposons que le paramètre à l'itération k soit donné par $\delta^{(k)} = h_2(\|f(x^{(k-m_k)})\|_2)$. Alors, on a*

$$\forall k \in \mathbb{N}, k - m_k \geq 1, \quad \|f(x^{(k)})\|_2 \leq C_2 \mu^{1/2} \|f(x^{(k-m_k-1)})\|_2^{3/2}.$$

Résultats numériques

1 Le problème

2 DIIS et accélération d'Anderson-Pulay

3 Analyse de convergence de variantes de DIIS

4 Résultats numériques

5 Et ensuite...

Data : $x^{(0)}$, tol, τ

$r^{(0)} = f(x^{(0)})$, $k = 0$, $m_k = 0$

while $\|r^{(k)}\|_2 > \text{tol}$ **do**

if $m_k = 0$ **then**

$x^{(k+1)} = g(x^{(k)})$

else

$\{\gamma_i^{(k)}\}_{i=1,\dots,m_k} = \arg \min_{(\gamma_1, \dots, \gamma_{m_k}) \in \mathbb{R}^{m_k}} \|r^{(k-m_k)} + \sum_{i=1}^{m_k} \gamma_i s^{(k-m_k+i)}\|_2$ (QR factorization)

$x^{(k+1)} = g(x^{(k-m_k)}) + \sum_{i=1}^{m_k} \gamma_i^{(k)} (g(x^{(k-m_k+i)}) - g(x^{(k-m_k)}))$

end

$r^{(k+1)} = f(x^{(k+1)})$ and $s^{(k+1)} = r^{(k+1)} - r^{(k-m_k)}$

 compute $\Pi_k s^{(k+1)}$ (orth. proj. of $s^{(k+1)}$ onto $\text{span}\{s^{(k-m_k+1)}, \dots, s^{(k)}\}$, given by the QR factorization)

if $\tau \|s^{(k+1)}\|_2 > \|(\text{id} - \Pi_k) s^{(k+1)}\|_2$ **then**

$m_{k+1} = 0$

else

$m_{k+1} = m_k + 1$

end

$k = k + 1$

end

Algorithme 1 : Variante avec redémarrage (R-CDIIS).

Data : $x^{(0)}$, tol , δ

$r^{(0)} = f(x^{(0)})$, $k = 0$, $m_k = 0$

while $\|r^{(k)}\|_2 > \text{tol}$ **do**

if $m_k = 0$ **then**

$x^{(k+1)} = g(x^{(k)})$

else

$\{\alpha_i^{(k)}\}_{i=1,\dots,m_k} = \arg \min_{(\alpha_1,\dots,\alpha_{m_k}) \in \mathbb{R}^{m_k}} \|r^{(k)} - \sum_{i=1}^{m_k} \alpha_i s^{(k-m_k+i)}\|_2$

$x^{(k+1)} = g(x^{(k)}) - \sum_{i=1}^{m_k} \alpha_i^{(k)} (g(x^{(k-m_k+i)}) - g(x^{(k-m_k+i-1)}))$

end

$r^{(k+1)} = f(x^{(k+1)})$

$s^{(k+1)} = r^{(k+1)} - r^{(k)}$

 set m_{k+1} the largest integer $m \leq m_k + 1$ such that for $k + 1 - m \leq i \leq k$, $\delta \|r^{(i)}\|_2 < \|r^{(k+1)}\|_2$

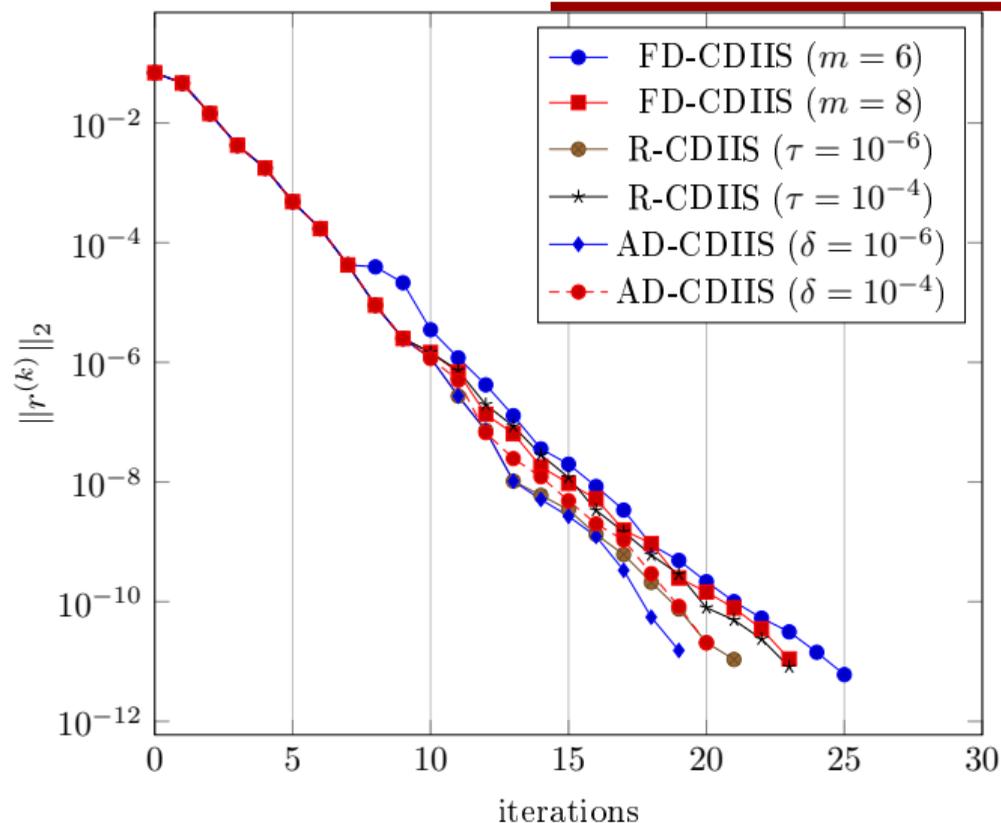
$k = k + 1$

end

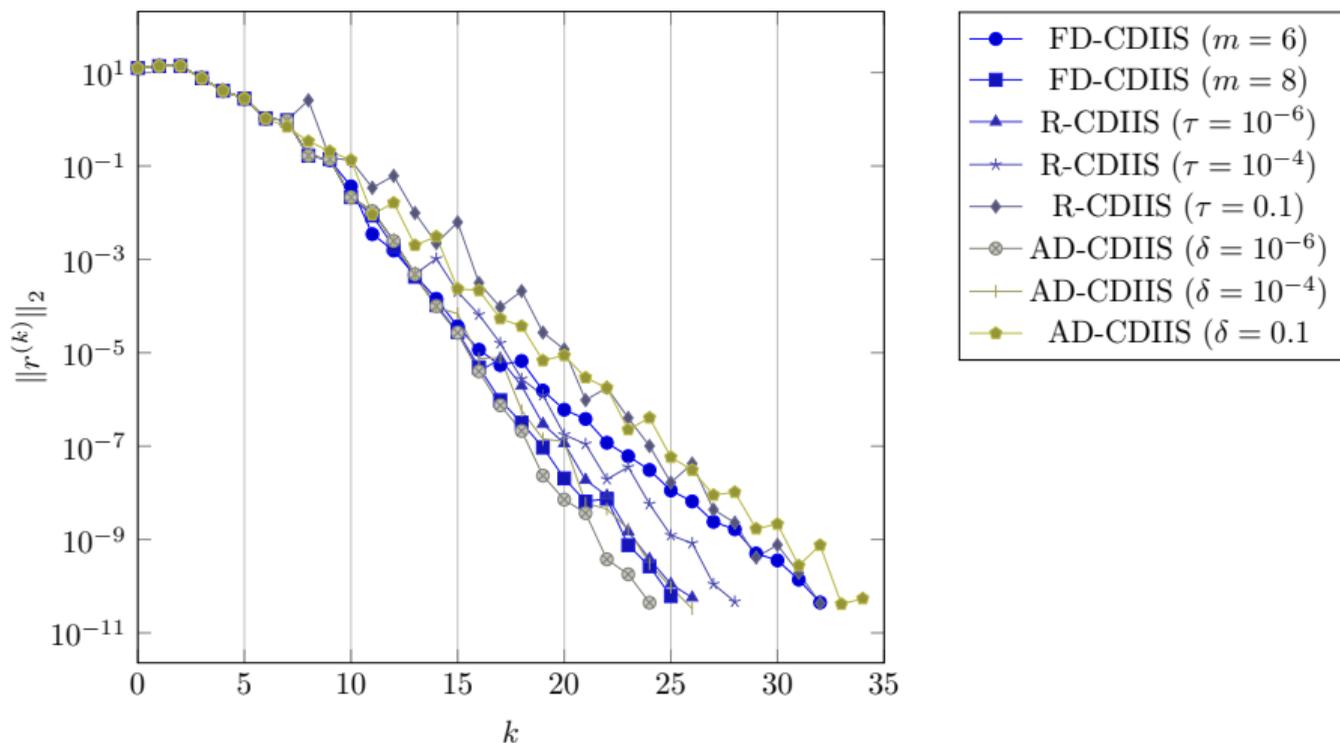
Algorithme 2 : Fenêtre adaptative (AD-CDIIS).

- Implémentation en Python avec PySCF, et code disponible

<https://plmlab.math.cnrs.fr/mchupin/restarted-and-adaptive-cdiis/>



Molécule Dimethylnitramine, modèle RHF, base 6-31G.



Molécule Cadmium-imidazole, modèle RKS/B3LYP, base 3-21G (condition initiale à partir de l'hamiltonien de cœur).

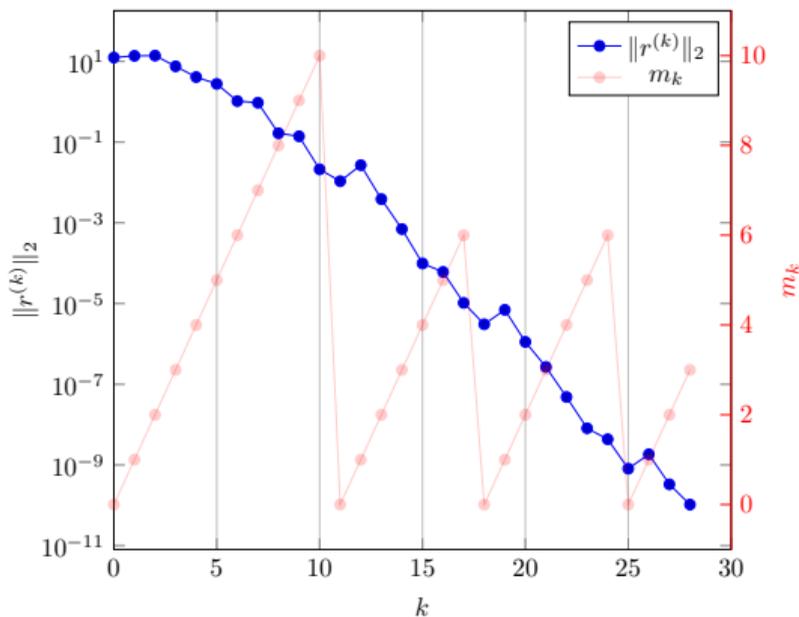
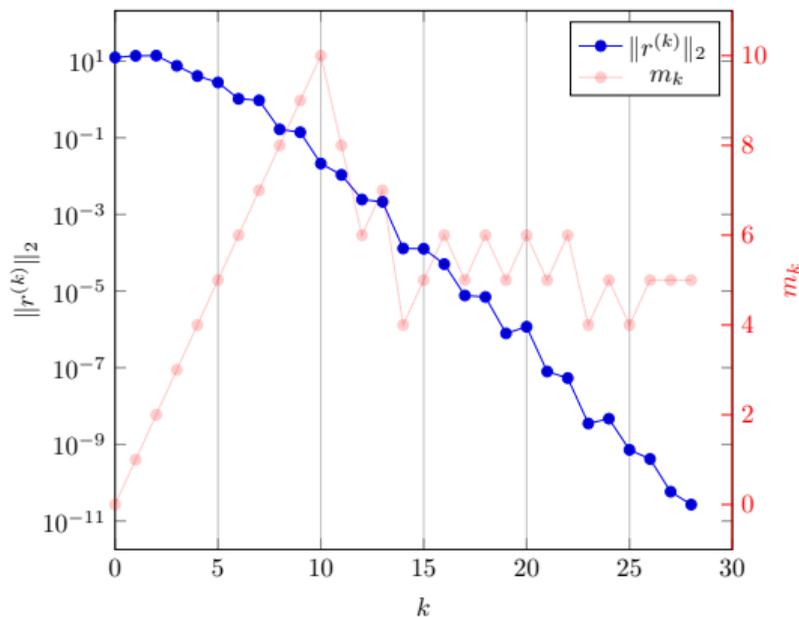
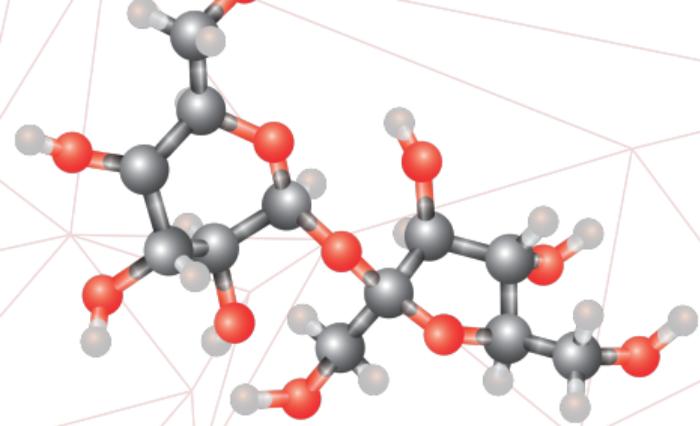
(a) *Restarted CDHIS* avec $\tau = 10^{-4}$.(b) *Adaptive-depth CDHIS* avec $\delta = 10^{-4}$.

Figure – Convergence de la norme du résidu et la profondeur de la fenêtre correspondante pour la molécule cadmium-imidazole.

- ▶ Diffuser ces algorithmes dans la communauté des chimistes (avec d'autres exemples, et *mix algorithmique*)
- ▶ Vérifier pour une autre modélisation (ROHF) ce qui est dit dans un papier récent [Benda, Vidal, Giner, Cancès, 2021]
- ▶ Étude de l'algorithme dans un autre contexte : cas linéaire et cas non linéaire (comparaison avec NLGMRES, etc.) [de Sterck *et al*, 2021]
- ▶ ...

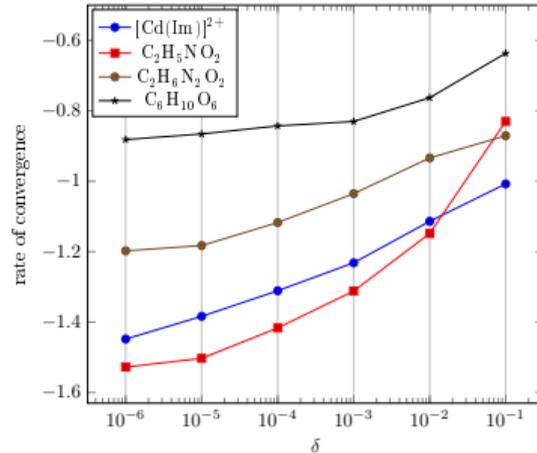
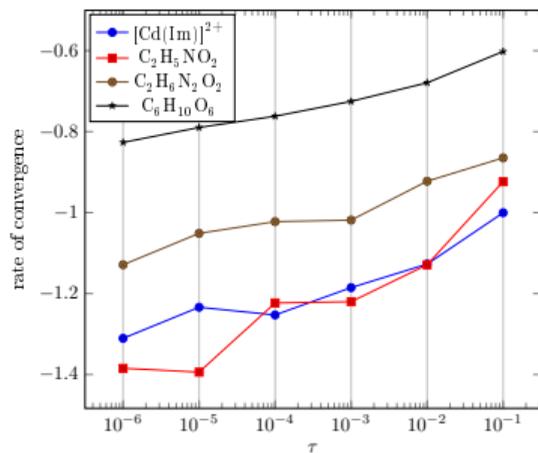
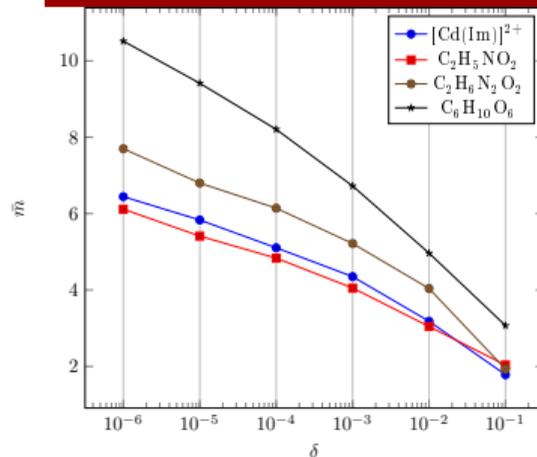
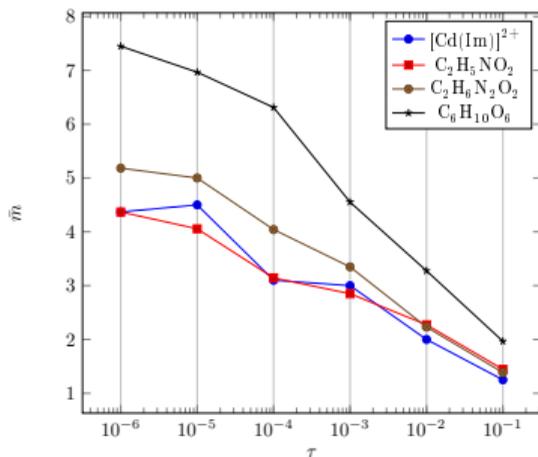
Publication

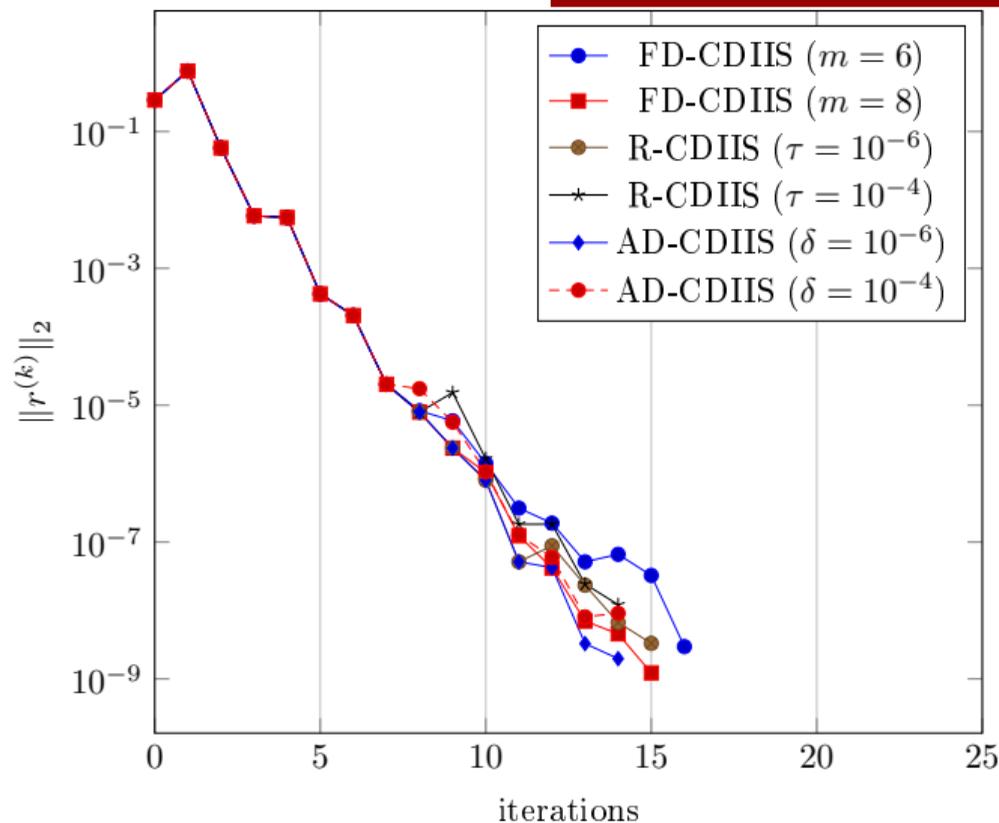
Convergence analysis of adaptive DIIS algorithms with application to electronic ground state calculations, Maxime CHUPIN, Mi-Song DUPUY, Guillaume LEGENDRE, Eric SÉRÉ, ESAIM : Mathematical Modelling and Numerical Analysis, EDP Sciences, 2021, 55 (6), pp.2785 - 2825. 10.1051/m2an/2021069



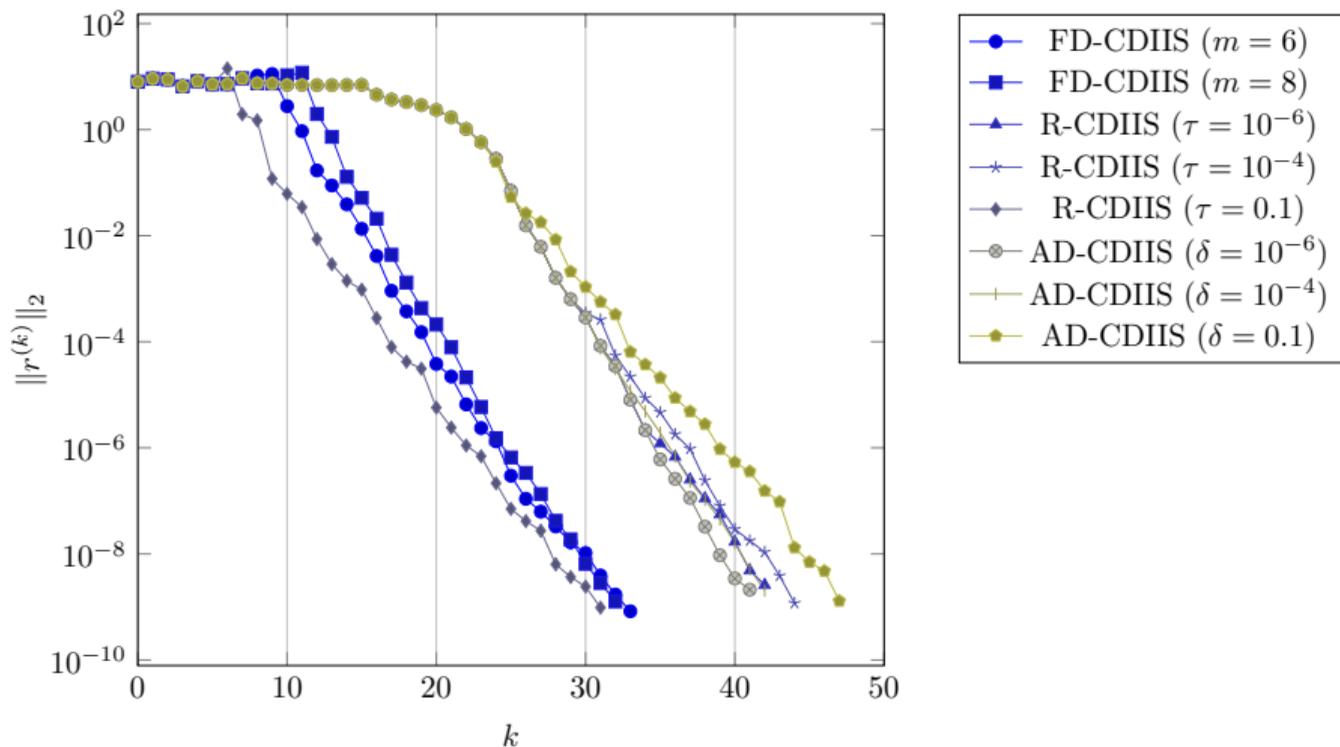
Merci pour votre attention!

Taille moyenne et taux de convergence





Molécule Cadmium-imidazole, modèle RKS/B3LYP, base 3-21G.



Molécule Dimethylnitramine, modèle RHF, base 6-31G (condition initiale à partir de l'hamiltonien de cœur).