

# Eficiencia computacional y ley de Amdahl para la técnica de simulación de resolución adaptativa

Luis Miguel Macedo Pinto

22 de mayo de 2024

## Resumen

Se compara la eficiencia computacional de la técnica de resolución adaptativa en simulación molecular con simulaciones atomísticas completas y de grano grueso equivalentes. Se muestra que la Ley de Amdahl adaptada puede estimar su eficiencia con precisión del 10

## 1. Introducción

La simulación de resolución adaptativa (AdResS) utiliza diferentes resoluciones moleculares según la región del sistema, lo que ahorra recursos computacionales. Se muestra que sigue la Ley de Amdahl en términos de eficiencia computacional. Se proporciona una fórmula general para estimar la ganancia computacional en comparación con simulaciones atomísticas completas. Se verifica numéricamente esta fórmula y se analiza la ganancia real de AdResS frente a simulaciones atomísticas completas.

En AdResS, el cuadro de simulación se divide en tres regiones con diferentes resoluciones moleculares: atomística, de grano grueso y mixta. Esta última contiene tanto resolución atomística como de grano grueso. La resolución en cada región está determinada por una función de ponderación, siendo una de las más comunes la que se utiliza en AdResS.

En AdResS, el sistema se divide en regiones con diferentes resoluciones moleculares, controladas por una función de ponderación. Las interacciones entre moléculas en diferentes regiones se interpolan mediante fuerzas. Un termostato global equilibra las transiciones de energía. La precisión se mejora con la introducción de una fuerza termodinámica, pero aumenta la carga computacional.

$$w(x) = \begin{cases} 1 & x < d_{AT} \\ \cos^2 \left[ \frac{\pi}{2(d_\Delta)} (x - d_{AT}) \right] & d_{AT} < x < d_{AT} + d_\Delta \text{ donde } d_{AT} \text{ y } d_\Delta \text{ son el} \\ 0 & d_{AT} + d_\Delta < x \end{cases}$$

Figura 1: Imagen formula

GC-AdResS extiende este enfoque a simulaciones tipo Gran Canónica, permitiendo un modelo de grano grueso más flexible. La eficiencia computacional varía dependiendo del tamaño de la región híbrida. En este artículo, no se aborda la fuerza termodinámica debido a su costo computacional.

$$F_{\alpha\beta} = w(X_\alpha) w(X_\beta) F_{\alpha\beta}^{atom} + [1 - w(X_\alpha) w(X_\beta)] F_{\alpha\beta}^{cm}$$

Figura 2: Imagen formula

## 2. Ley de Amdahl

La ley de Amdahl en informática predice la aceleración general de un proceso cuando solo una parte se mejora. En AdResS, aunque el cálculo de fuerza puede acelerarse, otras partes no. Utilizaremos esta idea para estimar la velocidad de AdResS en comparación con simulaciones atomísticas completas.

### 2.1. Análisis computacional

Para AdResS, no consideramos tiempos absolutos, sino que usamos el tiempo de una simulación atomística completa como referencia. El tiempo de ejecución en las simulaciones de Dinámica Molecular se divide en dos partes: el tiempo para calcular fuerzas (tF) y el tiempo para otras operaciones (tNF), como reconstrucción de lista de vecinos y comunicación, que escala con el número de átomos (N).

### 2.2. Simulaciones de grano grueso

La simulación de grano grueso representa el límite superior de aceleración, ya que una simulación AdResS con región atomística evanescente es equivalente a una simulación de grano grueso. En este contexto, no consideramos otros aspectos y nos enfocamos en la aceleración debido a la reducción del número de partículas. Esto implica un análisis del "peor de los casos".

## 3. Simulaciones AdResS

El límite superior en el rendimiento computacional de AdResS se mejora al considerar las zonas atomísticas, híbridas y de grano grueso. En la zona híbrida, se deben calcular interacciones de grano grueso y atomísticas. Esto se deduce asumiendo que el volumen es proporcional al número promedio de moléculas a una densidad dada.

$$t_{\text{AdResS}} = \frac{V_{\text{CG}} + V_{\text{HY}}}{V_{\text{Tot}}} t_{\text{CG}} + \frac{V_{\text{AT}} + V_{\text{HY}}}{V_{\text{Tot}}} t_{\text{AT}}.$$

Figura 3: Enter Caption

## 4. Implementaciones

En implementaciones como GROMACS y Espresso, la representación atomística no se elimina completamente en la región de grano grueso. Esto se hace para evitar operaciones costosas de reinserción de estructuras y para mantener la integración del grado de libertad interno. Esta decisión afecta al cálculo del factor de aceleración, corrigiendo el tiempo de cálculo.

$$t'_{\text{CG}} = t_{\text{F,AT}}/P^2 + t_{\text{NF,AT}}.$$

Figura 4: Descripción de la imagen.

## 5. Eficiencia/aceleración del rendimiento

Podemos generalizar las ecuaciones para sistemas AdResS con regiones atomísticas y híbridas de tamaño finito y definir el rendimiento computacional en comparación con simulaciones atomísticas. Al combinar las ecuaciones, obtenemos una fórmula integral.

$$t_{\text{AdResS}} = \frac{V_{\text{CG}} + V_{\text{HY}}}{V_{\text{Tot}}} (t_{\text{F,AT}}/P^2 + t_{\text{NF,AT}}) + \frac{V_{\text{AT}} + V_{\text{HY}}}{V_{\text{Tot}}} t_{\text{AT}} \quad (8)$$

and,

$$t_{\text{AdResS}}/t_{\text{AT}} = v + \left(1 + \frac{V_{\text{HY}}}{V_{\text{Tot}}} - v\right) (\%_{\text{F,AT}}/P^2 + 1 - \%_{\text{F,AT}}), \quad (9)$$

Figura 5: Descripción de la imagen.

## 6. Caso de mezclas

Para mezclas, el tiempo de cálculo tFAT se divide en contribuciones de las diferentes especies de moléculas. En una mezcla con dos tipos de moléculas A y B, tFAT se puede expresar en términos de contribuciones de diferentes tipos de interacciones.

$$\frac{t_{FAT}}{t_{AT}} = \%_{FAT,AA} + \%_{FAT,AB} + \%_{FAT,BB}$$

Figura 6: Descripción de la imagen.

Donde tAT es el tiempo de una simulación totalmente atomística. tAA es el tiempo utilizado en calcular fuerzas solo para interacciones A-A. De manera similar, se define tAB. No hay un método directo para calcular los porcentajes individuales, así que primero calculamos tAB directamente de la salida de GROMACS. Luego, calculamos tAA y consideramos el número de interacciones de tipos específicos en el sistema. Supongamos que hay X interacciones A-A, Y interacciones A-B y Z interacciones B-B.

## 7. Resultados numéricos

Comparamos la ganancia computacional predicha con la ganancia real de simulaciones AdResS versus simulaciones totalmente atomísticas. Analizamos cinco sistemas: agua, hexano, butano, DMSO y TBA/DMSO.

**Tabla 1**

Estimamos el tiempo necesario para calcular fuerzas en simulaciones atomísticas completas de varios sistemas. Sin embargo, la complejidad de la arquitectura computacional puede hacer que el tiempo real sea menor. Esta estimación teórica se considera una indicación del mejor rendimiento del cálculo computacional, dentro de un rango del 10

System	% <sub>FAT</sub>
TIP5P water	63
Butanol	58
Hexane	46
DMSO	55
TBA-DMSO mixture	57

Figura 7: Descripción de la imagen.

**Tabla 2**

Se compara la aceleración para agua TIP5P utilizando potencia IBI con presión rectificada en la región de grano grueso, calculada mediante Eq- y AdResS sin simulación. <sup>A</sup>T indica el tamaño de la región atomística. Las columnas reportan diferentes cantidades de la fórmula y la simulación, utilizando la misma convención en todas las tablas siguientes.

AT (nm)	$v$	Equation	Simulation
0.0	0.0	2.53	2.94
0.5	0.0136	2.45	2.69
0.9	0.0284	2.37	2.47
1.3	0.0511	2.26	2.17
1.7	0.0835	2.13	1.88
2.1	0.127	1.98	1.60

Figura 8: Descripción de la imagen.

**Tabla 3**

Aceleración real del agua IIPSP utilizando el potencial IBI con corrección de presión en la región de grano grueso. <sup>A</sup>T se refiere al tamaño de la región atomística.

AT (nm)	Actual speedup
0.0	2.45
0.5	2.24
0.9	2.05
1.3	1.81
1.7	1.56
2.1	1.33

Figura 9: Descripción de la imagen.

**Tabla 4**

Se compara la aceleración del butanol líquido usando el potencial IBI con corrección de presión en la región de grano grueso, calculado usando la ecuación 19, con la simulación AdResS.

La discrepancia entre la aceleración predicha y la obtenida de las simulaciones es del 15 por ciento, aunque en algunos casos, como el hexano, es menor al 10 por ciento. Una de las razones es el uso de potencial de grano grueso con corrección de presión, lo que afecta la densidad en la región híbrida y de grano grueso. Esto puede afectar el rendimiento ideal. esperado de AdResS. Aunque la discrepancia con respecto a la densidad es pequeña y no afecta las propiedades, podría influir en el rendimiento. La introducción de la fuerza termodinámica elimina el problema de la densidad no homogénea, pero implica un cálculo de

AT (nm)	$\nu$	Equation	Simulation
0.0	0.0	2.28	2.07
0.5	0.0127	2.22	1.98
1.0	0.0265	2.16	1.92
1.3	0.0476	2.08	1.82
1.7	0.0779	1.97	1.71
2.1	0.118	1.85	1.56

Figura 10: Descripción de la imagen.

fuerza adicional en la región híbrida. Aunque es pequeño, preferimos considerarlo en nuestra estimación del rendimiento. Los potenciales IBI son tabulados, lo que implica una penalización del 5 ciento. Además, para algunos sistemas, como TIP5P, es necesario calcular las interacciones de Coulomb, aunque esto no se ha considerado explícitamente. El sitio virtual adicional también añade una sobrecarga a la simulación AdResS en comparación con la atomística completa.

## 8. Conclusiones

Se desarrollo un modelo para predecir la máxima aceleración en simulaciones de resolución adaptativa, específicamente en AdResS. Confirmamos el modelo con precisión en múltiples sistemas. Para tamaños razonables de la región atomística, la ganancia puede ser significativa, lo que permite tratar sistemas más grandes con recursos computacionales estándar. Nuestra propuesta puede ser útil para que los investigadores estimen la ganancia al utilizar AdResS y decidan si es adecuado para su objetivo de ahorro de tiempo de cálculo.