

Notas de implementación para los códigos de los modelos coarse-grained

July 26, 2023

1 Consideraciones generales

1.1 Directiva de compilación

Se crea el ejecutable con una directiva como la que sigue:

```
gcc Main.c Read.c PDESolver.c Utilities.c -o EXE
```

Flags:

- Si se compila en Linux, añadir

`-lm`

- Para cuestiones de rendimiento, añadir

`-O2`

(el uso de

`-O3`

parece arriesgado).

1.2 Unidades

Por ahora, en los ficheros de entrada se entiende que las magnitudes están dadas en unidades del sistema internacional (metros y segundos **y, en principio, moles**). Internamente se trabaja de forma adimensional. Las salidas que dependen de espacio están dadas de forma adimensional. Otras salidas, según se especifique.

1.3 Ecuaciones

El modelo coarse-grained que trabajamos en la actualidad consta de

$$\frac{\partial c}{\partial t} = D_c \frac{\partial^2 c}{\partial x^2} - k_1 n(t, x) c - k_2 c + S, \quad (1.1)$$

$$\frac{\partial n}{\partial t} = D_n \frac{\partial^2 n}{\partial x^2} + \lambda_n(c) n. \quad (1.2)$$

Ambas ecuaciones se resuelven en un intervalo espacial. La primera puede tener condiciones de flujo cero o entrante, la segunda tiene siempre flujo cero. La primera ecuación podría no tener un término fuente extra en algunas aplicaciones. La ratio de proliferación coarse-grained $\lambda = \lambda_n(c)$ se calcula mediante

$$\frac{2F_S}{\tau_p} \frac{e^{-(\lambda+\nu)a_{G1/S}}}{\lambda + \nu + 1/\tau_p} = 1. \quad (1.3)$$

No se espera que este modelo dé un comportamiento razonable en escalas de tiempo cortas.

1.4 Formato de entrada

La simulación se ejecuta con una directiva de tipo

```
./EXE host_necrotic oxy_necrotic pars_pop pars_sim a
```

Por orden,

1. El primer argumento es el nombre del ejecutable.
2. El segundo y el tercero son las distribuciones iniciales de población y oxígeno.
3. El cuarto es un fichero de parámetros relativos a características de la población.
4. El quinto es un fichero de parámetros que indica información sobre la simulación.
5. El sexto es un literal de caracteres que identifica a los ficheros y carpetas de salida.

Formato para las distribuciones Ambos son ficheros dados por una única columna de valores. Leídos de arriba a abajo, se entiende que son los valores sobre una malla de paso uniforme (cuyo valor se especifica en otro fichero de entrada, ver más abajo), especificados de izquierda a derecha. Cada valor va separado del siguiente por un salto de línea. No se incluye un salto de línea tras el último valor (cuidado porque si se hace se procesa mal el tamaño de la malla espacial). El número de slots de la malla espacial se deduce de forma indirecta a partir del número de líneas de estos ficheros.

El fichero de entrada para la población se interpreta como una densidad de partículas (en cada localización espacial). El fichero de entrada para el oxígeno, de forma similar con valores de concentración. En la práctica, los valores en cuestión en cada slot de la malla se interpretan como moles por metro (slots de tamaño Δx , que *no tiene por qué* ser exactamente 1 metro).

Formato para el fichero de parámetros de la población Se incluyen una serie de números en la misma línea, separados por espacios. No hay ningún caracter tras el último valor numérico. Por orden, se trata de:

1. The death rate, ν .
2. The proliferation-related parameter, τ_p .
3. The value of a_+ , proliferation rate at zero oxygen. *De momento no vamos a considerar el régimen de $p6/p3$ para el cuál esta cantidad es relevante. Así que podríamos suprimirla.*
4. The value of $p6/p3$
5. The diffusion coefficient, D_n .
6. The survival rate (take $F_S = 1$ when we do not include therapy).

Formato para el fichero de parámetros de la simulación Se incluyen una serie de números en la misma línea, separados por espacios. No hay ningún caracter tras el último valor numérico. Por orden, se trata de:

1. El decay rate para el oxígeno, k_2 .
2. La tasa de consumo de oxígeno, k_1 .
3. El coeficiente de difusión del oxígeno, D_c .
4. El término fuente de la ecuación para el oxígeno, S .
5. El tiempo de parada de la simulación (suponiendo que empezamos a tiempo cero).
6. El número de ficheros de salida deseados para las distribuciones espaciales, $nfiles$ (aparte los de las condiciones iniciales). Equidistribuidos en tiempo entre el instante cero y el tiempo final especificado.
7. El paso de la malla, Δx .
8. El Δt sugerido para el integrador en tiempo (internamente se refina si fuera necesario para cumplir las condiciones de estabilidad pertinentes). Recomendable que divida de forma exacta al tiempo final.

Este formato puede variar (por el final) si hay que anhadir un flujo entrante, etc. Consultar el fichero de cabecera en cada caso. Todos los parámetros especificados van en unidades del SI siempre que sea aplicable.

1.5 Formato de salida

Se obtienen dos ficheros de salida y dos carpetas.

- Un fichero messages con información sobre los parámetros utilizados en la simulación asociada, expresados en las unidades del SI. También incluye los valores del equilibrio homogéneo no trivial asociado a los parámetros de la simulación proporcionados, expresado en las unidades del SI.
- Un fichero de dos columnas con el número total de individuos de la población (segunda columna, pueden ser decimales, realmente está expresado en moles) en función del tiempo (primera columna, en segundos), en los instantes de tiempo determinados conforme al tiempo final y el número de ficheros de salida especificados. Tales instantes están equiespaciados en tiempo y van correlativos a los ficheros en espacio que la rutina devuelve, ver siguiente punto.
- Una carpeta con $n_{\text{files}}+1$ ficheros que representan los valores de la densidad de población a cada instante de tiempo especificado (incluida la distribución inicial). El formato para cada uno de estos ficheros es el mismo que el de entrada para la condición inicial, **pero** estas salidas son valores adimensionales (para ello se dividen los valores dimensionales de la densidad de población por el valor del equilibrio no trivial de la población).
- Otra carpeta similar para los valores de la concentración de oxígeno en función del tiempo.

1.6 Sobre integradores en tiempo

De momento tenemos un Euler explícito, el más sencillo y de peores prestaciones. RK4 puede ser muy lento, ver si hay discrepancias que justifiquen su uso. Específicos de la GSL por ver si merece la pena, e implícitos lo mismo, ya que la ventaja de tomar pasos en tiempo grandes se puede ver anulada si esperamos variación temporal rápida del λ , o cuando los mezclamos con el híbrido ya que estamos subordinados al τ del Gillespie -otra cosa es que usemos algún tipo de tau-leaping, por investigar. Otro problema de los implícitos es que el modelo no es lineal, de forma que a cada paso hay que resolver un sistema grande de ecuaciones via Newton o similar, lo que puede ser bastante caro (al menos si se plantea de forma inocente) y anular la supuesta ventaja de tomar Δt grande.

2 Secuencia de pasos del código

2.1 Procesamiento de la entrada

Tras leer todos los ficheros de entrada de la simulación, se calculan los equilibrios no triviales (c_∞, n_∞) conforme a¹

$$c_\infty = c_{cr} + c_{cr} \left(\frac{1}{a-\nu} \log \left(\frac{2F_S}{1 + \nu\tau_p} \right) \right)^{-1/\beta}, \quad n_\infty = \frac{S - k_2 c_\infty}{k_1 c_\infty}.$$

Nótese que ahora mismo el valor $c_{cr} = c_{cr}(p^*)$ está dado por la expresión

$$c_{cr}(p^*) = \frac{c_{typ} c_{cr}(p^*)}{100 c_{cr}(1)}$$

donde c_{typ} es un valor típico para la concentración del oxígeno, y c_{cr} es la fórmula de Tomás para el oxígeno crítico.

Esta información (c_∞, n_∞) se escribe en el fichero de messages. Se tomarán además un tiempo y un espacio adimensionales dados por

$$\hat{t} = t/\tau_p, \quad \hat{x} = x/l,$$

donde la escala característica en longitud viene dada por

$$l = \sqrt{D_n \tau_p}.$$

Cuestión importante: El valor de c_{typ} ahora mismo es una constante inventada que supondré que estoy midiendo en moles por metro. Adoptar estas unidades arrastra al resto del desarrollo del código, como paso a explicar. Si los inconvenientes son importantes podemos probar a adoptar otras unidades (nanomolar o lo que nos parezca oportuno), ya que por ahora todo es ficticio.

Adoptar unidades de moles por metro para el oxígeno típico hace que sea natural que midamos las concentraciones de entrada de oxígeno en valores de moles por metro. De modo que n_∞ también es natural medirla en moles por metro y así también el fichero de entrada de la población (luego internamente se adimensionaliza todo). Hay que vigilar que los valores de los equilibrios para la población no salgan anormalmente grandes. En cualquier caso, en este estadio no creo que podamos aspirar a comparaciones con datos reales.

2.2 Cálculos iniciales

Se sobreentiende que de aquí en adelante trabajamos con

$$\hat{c}(\hat{t}, \hat{x}) = c(t, x)/c_\infty, \quad \hat{n}(\hat{t}, \hat{x}) = n(t, x)/n_\infty,$$

¹No sé si hay alguna condición cómoda para poder asegurar a priori que $n_\infty, c_\infty \geq 0$.

donde c y n son las soluciones de (1.1)–(1.2). Se lleva a cabo esta operación con las lecturas iniciales de población y oxígeno.

Se definen los parámetros adimensionales

$$\hat{D}_c = D_c/D_n, \quad \hat{k}_1 = k_1 n_\infty \tau_p, \quad \hat{k}_2 = k_2 \tau_p, \quad \hat{S} = S \tau_p / c_\infty, \quad \hat{\nu} = \nu \tau_p.$$

Se definen además

$$\hat{a}_{G_1 \setminus S}(\hat{c}) = a_{G_1 \setminus S}(c)/\tau_p, \quad \hat{\lambda}(\hat{c}) = \lambda(c).$$

Se calculan las edades iniciales de división, basándonos en el modelo (dimensional)

$$a_{G_1 \setminus S}(c) = a_- \left(\frac{c}{c_{cr}(p^*)} - 1 \right)^{-\beta}.$$

Aquí $p^* = p6/p3$ que de momento se está tomando como igual a uno. Para realizar los cálculos adimensionales, se introduce $\hat{a} = a/\tau_p$ y por tanto se reescribe todo como

$$\hat{a}_{G_1 \setminus S}(\hat{c}) = \frac{a_{G_1 \setminus S}(c)}{\tau_p} = \hat{a}_- \left(\frac{\hat{c}}{\hat{c}_{cr}(p^*)} - 1 \right)^{-\beta}$$

El valor a_- está fijado a 8.5 horas (30600 segundos), al menos por ahora.

2.3 Bucle principal

Se simula una versión adimensional de las ecuaciones. Conforme a todo lo descrito anteriormente, se trata de

$$\frac{\partial \hat{c}}{\partial \hat{t}} = \hat{D}_c \frac{\partial^2 \hat{c}}{\partial \hat{x}^2} - \hat{k}_1 \hat{n} \hat{c} - \hat{k}_2 \hat{c} + \hat{S}, \quad (2.1)$$

$$\frac{\partial \hat{n}}{\partial \hat{t}} = \frac{\partial^2 \hat{n}}{\partial \hat{x}^2} + \hat{\lambda}_n(\hat{c}) \hat{n}. \quad (2.2)$$

Para discretizar las derivadas espaciales se usa la fórmula centrada de orden 2, con respecto al paso de malla adimensionalizado $\Delta \hat{x} = \Delta x/l$. A este fin, internamente se consideran coeficientes de difusión efectivos dados por

$$\bar{D}_c = \frac{\hat{D}_c}{(\Delta \hat{x})^2}, \quad \bar{D}_n = \frac{1}{(\Delta \hat{x})^2}.$$

A cada paso del bucle principal se intenta avanzar el sistema un incremento de tiempo de Δt segundos; en el modelo adimensionalizado (2.1)-(2.2), esto se traduce en realizar un avance de $\Delta \hat{t} = \Delta t/\tau_p$. Si este incremento de tiempo no cumple las condiciones de estabilidad para hacer un Euler explícito con (2.1)-(2.2), entonces se refina internamente tal incremento de forma que las cumpla.

Para realizar el cálculo de la tasa neta de proliferación (adimensionalizada) que va en (2.2) se reescribe (1.3) como

$$\frac{1}{2F_S} = \frac{\exp\left(-\hat{a}_{G_1 \setminus S}(\hat{c})(\hat{\nu} + \hat{\lambda}(\hat{c}))\right)}{1 + \hat{\nu} + \hat{\lambda}(\hat{c})}.$$

Que a su vez se resuelve tomando logaritmos, es decir, se le aplica Newton-Rhapson a

$$\log(1 + \hat{\nu} + \hat{\lambda}(\hat{c})) - \log(2F_S) + \hat{a}_{G_1 \setminus S}(\hat{c})(\hat{\nu} + \hat{\lambda}(\hat{c})) = 0.$$

Tras hacer un avance de $\Delta\hat{t}$ se vuelven a calcular los umbrales de división tal como se explicó anteriormente.

2.4 Datos de salida y terminación

Se divide el intervalo temporal $[0, tstop]$ en n files subintervalos de igual longitud. A cada nodo temporal así obtenido se mandan datos a la salida, conforme se explicó en la subsección 1.5. Como los datos espaciales se muestran adimensionales, no es necesaria ninguna reconversión previa a la salida. Para el fichero del “total de células” se multiplican los valores adimensionales de la población en cada casilla por $n_\infty \Delta x$ (modelo genuinamente 1d) y se suman sobre todas las casillas. **Este valor está expresado en moles.** E.g. si por entrada pasamos $\Delta x = 10^{-3}$, se entiende que el paso de malla es un milímetro y el número total de células se obtiene deshaciendo la normalización de la densidad por el valor de equilibrio y multiplicando por el tamaño en metros de cada caja, luego sumando las contribuciones de todas las cajas.

Nótese que si la simulación aborta antes de terminar, sí se generan las salidas espaciales previas al instante de fallo, pero no el fichero de mensajes ni el de número total de células vs tiempo.