Fat vs. thin threading approach on GPUs application to stochastic simulation of chemical reactions

Supplemental material

Guido Klingbeil, Radek Erban, Mike Giles, and Philip K. Maini

This document presents the six chemical reaction systems used to compare the performance of the fat and thin threading approaches.

1 Dimerisation-decay model

The decay-dimerisation reaction system is taken from Gillespie [1]. It consists of 3 molecular species and 4 reactions. The monomer X_1 decays or dimerises reversibly into the dimer X_2. X_2 itself is unstable and turns into a stable form X_3:

$$\begin{aligned} R_1: &\quad X_1 \xrightarrow{c_1} \emptyset \\ R_2: &\quad 2X_1 \xrightarrow{c_2} X_2 \\ R_3: &\quad X_2 \xrightarrow{c_3} 2X_1 \\ R_4: &\quad X_2 \xrightarrow{c_4} X_3 \end{aligned}$$

The chosen reaction rates are $c_1 = 1\,\mathrm{h}^{-1}, c_2 = 0.002\,\mathrm{h}^{-1}, c_3 = 0.5\,\mathrm{h}^{-1}, c_4 = 0.04\,\mathrm{h}^{-1}$ and the initial molecular populations at time $t = 0\,\mathrm{sec}$ are $X_1 = 100000, X_2 = X_3 = 0$.

2 Oregonator

The Oregonator is a model simulating the oscillating Belousov-Zhabotinskii reaction developed by Field, Körös and Noyes [2]. The reaction system of the Oregonator describing the

general kinetic scheme of the Belousov-Zhabotinskii reaction is given by:

$$
\begin{aligned}
R_1: &\quad \bar{X}_1 + Y_2 \xrightarrow{c_1} Y_1 \\
R_2: &\quad Y_1 + Y_2 \xrightarrow{c_2} Z_1 \\
R_3: &\quad \bar{X}_2 + Y_1 \xrightarrow{c_3} 2\ Y_1 + Y_3 \\
R_4: &\quad 2Y_1 \xrightarrow{c_4} Z_2 \\
R_5: &\quad \bar{X}_3 + Y_3 \xrightarrow{c_5} Y_2
\end{aligned}
$$

The species marked with a bar are assumed to be constant. The reaction rates are $c_1 = 10.0\,\mathrm{h}^{-1}$, $c_2 = 0.5\,\mathrm{h}^{-1}$, $c_3 = 104.0\,\mathrm{h}^{-1}$, $c_4 = 0.016\,\mathrm{h}^{-1}$, $c_5 = 1.04\,\mathrm{h}^{-1}$ and the initial conditions at $t = 0\,\mathrm{sec}$ for the molecular populations are $Y_1 = 500$, $Y_2 = 1000$, $Y_3 = 2000$, $Z_1 = 2000$, $Z_2 = 50000$ [3].

3 Circadian cycle

The circadian rhythm is an approximately 24-hour cycle in biochemical or behavioural processes of many living entities, including plants, animals, and bacteria. The model used is the simplified circadian cycle model by Vilar *et al.* [4] based on the model of Barkai and Leibler [5].

The biochemical network of the circadian oscillator model is given in Figure 1. The core of the network is intracellular transcription regulation of the two genes involved, an activator gene D_A and a repressor gene D_R. Both are transcribed into mRNA M_A and M_R, respectively, and subsequently translated into the activator protein A and repressor protein R. The activator A binds to the A and R promoters simultaneously, increasing their transcription. A acts as the positive element in transcription, whereas R acts as the negative element by repressing the activator. The cycle is completed by repressor degradation and re-expression of the activator [4].

The molecular species of the circadian cycle model are:

- activator DNA D_A
- activator mRNA M_A
- activator protein A
- activator DNA-promoter complex D'_A
- repressor DNA D_R
- repressor mRNA M_R
- repressor protein R
- repressor DNA-promoter complex D'_R
- inactivated activator-repressor complex C

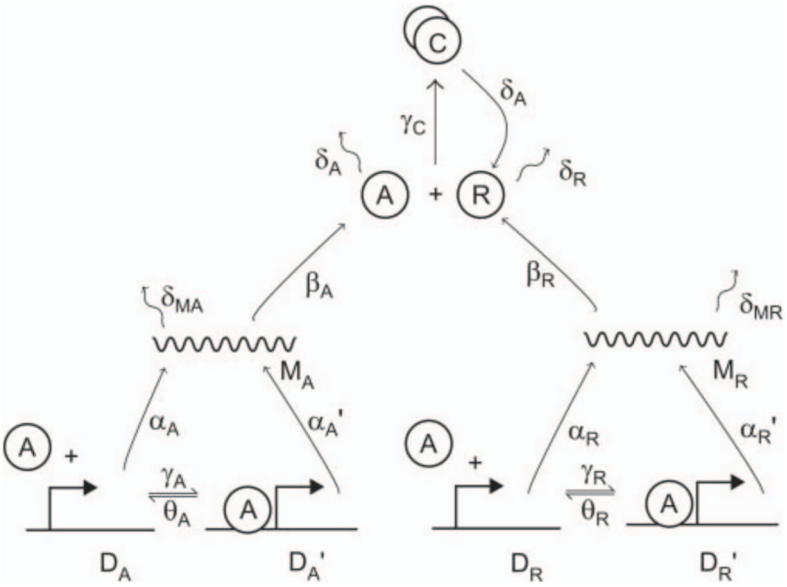

Figure 1: Biochemical network of the circadian oscillator reaction system. The core of the network is intracellular transcription regulation of the two genes involved, an activator gene D_A and a repressor gene D_R. Oscillations arise since the activator binds to the promoters of both genes simultaneously. Thus with the activator A the repressor R is expressed. The repressor R, in turn, inactivates A forming the complex C. Taken with permission from Vilar *et al.* [4].

The reactions of the circadian cycle model are:

$$
\begin{aligned}
R_1 &: \quad A + D_A \xrightarrow{\gamma_A} D'_A \\
R_2 &: \quad D'_A \xrightarrow{\Theta_A} D_A + A \\
R_3 &: \quad A + D_R \xrightarrow{\gamma_R} D'_R \\
R_4 &: \quad D'_R \xrightarrow{\Theta_R} D_R + A \\
R_5 &: \quad D_A \xrightarrow{\alpha_A} M_A \\
R_6 &: \quad D'_A \xrightarrow{\alpha'_A} M_A \\
R_7 &: \quad D_R \xrightarrow{\alpha_R} M_R \\
R_8 &: \quad D'_R \xrightarrow{\alpha'_R} M_R \\
R_9 &: \quad M_A \xrightarrow{\delta_{M_A}} \emptyset \\
R_{10} &: \quad M_A \xrightarrow{\beta_A} A \\
R_{11} &: \quad M_R \xrightarrow{\delta_{M_R}} \emptyset
\end{aligned}
$$

$$
\begin{aligned}
R_{12}: &\quad M_R \xrightarrow{\beta_R} R \\
R_{13}: &\quad A + R \xrightarrow{\gamma_C} C \\
R_{14}: &\quad A \xrightarrow{\delta_A} \emptyset \\
R_{15}: &\quad R \xrightarrow{\delta_R} \emptyset \\
R_{16}: &\quad C \xrightarrow{\delta_A} R
\end{aligned}
$$

with the reaction rates $\alpha_A = 50\,\mathrm{h}^{-1}$, $\alpha'_A = 500\,\mathrm{h}^{-1}$, $\alpha_R = 0.01\,\mathrm{h}^{-1}$, $\alpha'_R = 50\,\mathrm{h}^{-1}$, $\beta_A = 50\,\mathrm{h}^{-1}$, $\beta_R = 5\,\mathrm{h}^{-1}$, $\delta_{M_A} = 10\,\mathrm{h}^{-1}$, $\delta_{M_R} = 0.5\,\mathrm{h}^{-1}$, $\delta_A = 1\,\mathrm{h}^{-1}$, $\delta_R = 0.2\,\mathrm{h}^{-1}$, $\gamma_A = 1\,\mathrm{h}^{-1}$, $\gamma_R = 1\,\mathrm{h}^{-1}$, $\gamma_C = 2\,\mathrm{h}^{-1}$, $\Theta_A = 100\,\mathrm{h}^{-1}$. The initial number of molecules at $t = 0\,\mathrm{sec}$ are $D_A = D_R = 1$, $D'_A = D'_R = M_A = M_R = A = R = C = 0$. Since the model assumes the complex C turns into R by degradation of A, the rate δ_A appears twice [4].

4 *lac*-operon

This simplified model of the *lac*-operon is taken from Wilkinson [6]. The *lac*-operon is composed of a promoter P, the operator Op and three genes $lacZ$, $lacY$, and $lacA$. Of these three genes, only the $lacZ$ expressing β-galactosidase is part of the model. β-galactosidase is an intracellular enzyme cleaving the disaccharide lactose into glucose and galactose. The inhibitor I binds either to lactose L or the operator Op. If the inhibitor is bound to the operon, its transcription is prevented. Thus in the presence of lactose fewer inhibitor molecules bind to the operon and the operon's expression level increases [7].
The molecular species of the *lac*-operon model are:

- inhibitor gene I_{DNA}
- inhibitor transcript I_{RNA}
- Inhibitor protein I
- operon Op
- RNA polymerase $RNAp$
- RNA polymerase bound to operon $RNAp$-Op
- operon transcript RNA
- β-galactosidase Z
- lactose L
- lactose bound to inhibitor I-L
- operon bound to inhibitor I-Op

The reactions of the *lac*-operon model are:

$$
\begin{aligned}
R_1: &\quad I_{DNA} \xrightarrow{c_1} I_{DNA} + I_{RNA} \\
R_2: &\quad I_{RNA} \xrightarrow{c_2} I_{RNA} + I \\
R_3: &\quad I + L \xrightarrow{c_3} I\text{-}L \\
R_4: &\quad I\text{-}L \xrightarrow{c_4} I + L \\
R_5: &\quad I + Op \xrightarrow{c_5} I\text{-}Op \\
R_6: &\quad I\text{-}Op \xrightarrow{c_6} I + Op \\
R_7: &\quad Op + RNAp \xrightarrow{c_7} RNAp\text{-}Op \\
R_8: &\quad RNAp\text{-}Op \xrightarrow{c_8} Op + RNAp \\
R_9: &\quad RNAp\text{-}Op \xrightarrow{c_9} Op + RNAp + RNA \\
R_{10}: &\quad RNA \xrightarrow{c_{10}} RNA + Z \\
R_{11}: &\quad L + Z \xrightarrow{c_{11}} Z \\
R_{12}: &\quad I_{RNA} \xrightarrow{c_{12}} \emptyset \\
R_{13}: &\quad I \xrightarrow{c_{13}} \emptyset \\
R_{14}: &\quad I\text{-}L \xrightarrow{c_{14}} L \\
R_{15}: &\quad RNA \xrightarrow{c_{15}} \emptyset \\
R_{16}: &\quad Z \xrightarrow{c_16} \emptyset
\end{aligned}
$$

The reaction rates are $c_1 = 0.02\,\mathrm{h}^{-1}$, $c_2 = 0.1\,\mathrm{h}^{-1}$, $c_3 = 0.005\,\mathrm{h}^{-1}$, $c_4 = 0.1\,\mathrm{h}^{-1}$, $c_5 = 1\,\mathrm{h}^{-1}$, $c_6 = 0.01\,\mathrm{h}^{-1}$, $c_7 = 0.1\,\mathrm{h}^{-1}$, $c_8 = 0.01\,\mathrm{h}^{-1}$, $c_9 = 0.03\,\mathrm{h}^{-1}$, $c_{10} = 0.1\,\mathrm{h}^{-1}$, $c_{11} = 1e-5\,\mathrm{h}^{-1}$, $c_{12} = 0.01\,\mathrm{h}^{-1}$, $c_{13} = 0.002\,\mathrm{h}^{-1}$, $c_{14} = 0.002\,\mathrm{h}^{-1}$, $c_{15} = 0.01\,\mathrm{h}^{-1}$, $c_{16} = 0.001\,\mathrm{h}^{-1}$. The chosen initial molecular populations at $t = 0\,\mathrm{sec}$ are $I_{DNA} = 10$, $I_{RNA} = 0$, $I = 50$, $Op = 10$, $RNAp = 1000$, $RNA = 0$, $Z = 0$, $L = 1640000$, $I\text{-}L = 0$, $I\text{-}Op = 0$, $RNAp\text{-}Op = 0$. The number of RNA polymerases $RNAp$ is kept constant [6].

5 Fully connected reaction network

The fully connected reaction network consists of 6 chemical species X_1 to X_6 which can be reversibly converted into each other at a reaction rate of $c = 1\,\mathrm{h}^{-1}$. Initially all molecules are of species X_1.
The fully connected reaction network consists of 6 species and 30 reactions:

$$
\begin{aligned}
R_1:\ X_1 &\xrightarrow{c} X_2\ , & R_2:\ X_2 &\xrightarrow{c} X_1 \\
R_3:\ X_1 &\xrightarrow{c} X_4\ , & R_4:\ X_4 &\xrightarrow{c} X_1 \\
R_5:\ X_1 &\xrightarrow{c} X_5\ , & R_6:\ X_5 &\xrightarrow{c} X_1 \\
R_7:\ X_1 &\xrightarrow{c} X_3\ , & R_8:\ X_3 &\xrightarrow{c} X_1
\end{aligned}
$$

$$
\begin{array}{llll}
R_9: & X_1 \xrightarrow{c} X_6\,, & R_{10}: & X_6 \xrightarrow{c} X_1 \\
R_{11}: & X_2 \xrightarrow{c} X_3\,, & R_{12}: & X_3 \xrightarrow{c} X_2 \\
R_{13}: & X_2 \xrightarrow{c} X_4\,, & R_{14}: & X_4 \xrightarrow{c} X_2 \\
R_{15}: & X_2 \xrightarrow{c} X_5\,, & R_{16}: & X_5 \xrightarrow{c} X_2 \\
R_{17}: & X_2 \xrightarrow{c} X_6\,, & R_{18}: & X_6 \xrightarrow{c} X_2 \\
R_{19}: & X_3 \xrightarrow{c} X_4\,, & R_{20}: & X_2 \xrightarrow{c} X_3 \\
R_{21}: & X_3 \xrightarrow{c} X_5\,, & R_{22}: & X_5 \xrightarrow{c} X_3 \\
R_{23}: & X_3 \xrightarrow{c} X_6\,, & R_{24}: & X_6 \xrightarrow{c} X_3 \\
R_{25}: & X_4 \xrightarrow{c} X_5\,, & R_{26}: & X_5 \xrightarrow{c} X_4 \\
R_{27}: & X_4 \xrightarrow{c} X_6\,, & R_{28}: & X_6 \xrightarrow{c} X_4 \\
R_{29}: & X_5 \xrightarrow{c} X_6\,, & R_{30}: & X_6 \xrightarrow{c} X_5
\end{array}
$$

The initial molecular populations at time $t = 0\,\mathrm{sec}$ are $X_1 = 1000000$, $X_2 = X_3 = X_4 = X_5 = X_6 = 0$.

6 Multiple dimerisation-decay models

The multiple dimerisation-decay reaction system is based upon the reaction system introduced in Section 1. Multiple dimerisation-decay models merges 2, 4 and 6 decay-dimerisation reactions systems into one larger reaction system. Each of the dimerisation-decay models has an independent set of molecular species and reactions.
8 reactions, 6 species system:

$$
\begin{array}{ll}
R_1: & X_1 \xrightarrow{c_1} \emptyset \\
R_2: & 2X_1 \xrightarrow{c_2} X_2 \\
R_3: & X_2 \xrightarrow{c_3} 2X_1 \\
R_4: & X_2 \xrightarrow{c_4} X_3 \\
\hline
R_5: & X_4 \xrightarrow{c_1} \emptyset\,, \\
R_6: & 2X_4 \xrightarrow{c_2} X_5 \\
R_7: & X_5 \xrightarrow{c_3} 2X_4 \\
R_8: & X_5 \xrightarrow{c_4} X_6
\end{array}
$$

16 reactions, 12 species system:

$$
\begin{array}{ll}
R_1: & X_1 \xrightarrow{c_1} \emptyset \\
R_2: & 2X_1 \xrightarrow{c_2} X_2
\end{array}
$$

$$
\begin{aligned}
R_3 &: \quad X_2 \xrightarrow{c_3} 2X_1 \\
R_4 &: \quad X_2 \xrightarrow{c_4} X_3
\end{aligned}
$$

$$
\begin{aligned}
R_5 &: \quad X_4 \xrightarrow{c_1} \emptyset\ , \\
R_6 &: \quad 2X_4 \xrightarrow{c_2} X_5 \\
R_7 &: \quad X_5 \xrightarrow{c_3} 2X_4 \\
R_8 &: \quad X_5 \xrightarrow{c_4} X_6
\end{aligned}
$$

$$
\begin{aligned}
R_9 &: \quad X_7 \xrightarrow{c_1} \emptyset \\
R_{10} &: \quad 2X_7 \xrightarrow{c_2} X_8 \\
R_{11} &: \quad X_8 \xrightarrow{c_3} 2X_7 \\
R_{12} &: \quad X_8 \xrightarrow{c_4} X_9
\end{aligned}
$$

$$
\begin{aligned}
R_{13} &: \quad X_{10} \xrightarrow{c_1} \emptyset \\
R_{14} &: \quad 2X_{10} \xrightarrow{c_2} X_{11} \\
R_{15} &: \quad X_{11} \xrightarrow{c_3} 2X_{10} \\
R_{16} &: \quad X_{11} \xrightarrow{c_4} X_{12}
\end{aligned}
$$

24 reactions, 18 species system:

$$
\begin{aligned}
R_1 &: \quad X_1 \xrightarrow{c_1} \emptyset \\
R_2 &: \quad 2X_1 \xrightarrow{c_2} X_2 \\
R_3 &: \quad X_2 \xrightarrow{c_3} 2X_1 \\
R_4 &: \quad X_2 \xrightarrow{c_4} X_3
\end{aligned}
$$

$$
\begin{aligned}
R_5 &: \quad X_4 \xrightarrow{c_1} \emptyset \\
R_6 &: \quad X_4 \xrightarrow{c_2} X_5 \qquad 2 \\
R_7 &: \quad X_5 \xrightarrow{c_3} 2X_4 \\
R_8 &: \quad X_5 \xrightarrow{c_4} X_6
\end{aligned}
$$

$$
\begin{aligned}
R_9 &: \quad X_7 \xrightarrow{c_1} \emptyset \\
R_{10} &: \quad 2X_7 \xrightarrow{c_2} X_8 \\
R_{11} &: \quad X_8 \xrightarrow{c_3} 2X_7 \\
R_{12} &: \quad X_8 \xrightarrow{c_4} X_9
\end{aligned}
$$

$$
\begin{aligned}
R_{13}: &\quad X_{10} \xrightarrow{c_1} \emptyset \\
R_{14}: &\quad 2X_{10} \xrightarrow{c_2} X_{11} \\
R_{15}: &\quad X_{11} \xrightarrow{c_3} 2X_{10} \\
R_{16}: &\quad X_{11} \xrightarrow{c_4} X_{12}
\end{aligned}
$$

$$
\begin{aligned}
R_{17}: &\quad X_{13} \xrightarrow{c_1} \emptyset \\
R_{18}: &\quad 2X_{13} \xrightarrow{c_2} X_{14} \\
R_{19}: &\quad X_{14} \xrightarrow{c_3} 2X_{13} \\
R_{20}: &\quad X_{14} \xrightarrow{c_4} X_{15}
\end{aligned}
$$

$$
\begin{aligned}
R_{21}: &\quad X_{16} \xrightarrow{c_1} \emptyset \\
R_{22}: &\quad 2X_{16} \xrightarrow{c_2} X_{17} \\
R_{23}: &\quad X_{17} \xrightarrow{c_3} 2X_{16} \\
R_{24}: &\quad X_{17} \xrightarrow{c_4} X_{18}
\end{aligned}
$$

The chosen reaction rates are $c_1 = 1\,\mathrm{h}^{-1}, c_2 = 0.002\,\mathrm{h}^{-1}, c_3 = 0.5\,\mathrm{h}^{-1}, c_4 = 0.04\,\mathrm{h}^{-1}$ and the initial molecular populations at time $t = 0\,\mathrm{sec}$ are $X_1 = X_4 = X_7 = X_{10} = X_{13} = X_{16} = 100000$. All other initial molecular populations are 0.

References

[1] D. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems," *Journal of Chemical Physics*, vol. 115, no. 4, pp. 1716–1733, 2001.

[2] R. Field and R. Noyes, "Oscillations in chemical systems. iv. limit cycle behavior in a model of a real chemical reaction," *The Journal of Chemical Physics*, vol. 60, no. 5, pp. 1877–1884, March 1974.

[3] D. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *Journal of Physical Chemistry*, vol. 81, no. 25, pp. 2340–2361, 1977.

[4] J. Vilar, H. Kueh, N. Barkai, and S. Leibler, "Mechanisms of noise-resistance in genetic oscillators," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 9, pp. 5988–5992, 2002.

[5] M. Barkai and S. Leibler, "Biological rhythms: Circadian clocks limited by noise," *Nature*, vol. 403, pp. 267–268, 2000.

[6] D. Wilkinson, *Stochastic Modelling for Systems Biology*. Chapman & Hall/CRC, 2006.

[7] R. Knippers, *Molekulare Genetik*, 9th ed. Stuttgart: Georg Thieme Verlag, 2006.