

- [16] Martin, W.R.W., Palmer, M.R., Patlak, C.S., Calne, D.B.: Nigrostriatal Function in Humans Studied with Positron Emission Tomography. *Ann. Neurol.* 26, 535–542 (1989)
- [17] Melega, W.P., Hoffman, J.M., Luxen, A., Nissenson, C.H., Phelps, M.E., Barrio, J.R.: The effects of carbidopa on the metabolism of 6-[¹⁸F]fluorodopa-L-DOPA in rats, monkeys and humans. *Life Sci.* 47, 149–157 (1990)
- [18] Patlak, C.S., Blasberg, R.G.: Graphical Evaluation of Blood-to-Brain Transfer Constants from Multiple-Time uptake data. Generalizations. *J. Cereb. Blood Flow Metab.* 5, 584–590 (1985)
- [19] Patlak, C.S., Blasberg, R.G., Fenstermacher, J.D.: Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J. Cereb. Blood Flow Metab.* 3, 1–7 (1983)
- [20] Rousset, O.G., Deep, P., Kuwabara, H., Evans, A.C., Gjedde, A.H., Cumming, P.: Effect of partial volume correction on estimates of the influx and cerebral metabolism of 6-[¹⁸F]fluoro-L-dopa Studied with PET in normal control and Parkinson's Disease subjects. *Synapse* 37, 81–89 (2000)

Supplementary Material

Hereby, we show $Y_2(t)$, which is not shown in Eq. (11), at the following URL:

<http://www.math.kyushu-u.ac.jp/~phirosi/pet/Y2.pdf>,

where, for instance, $Root[k_1\#1 + k_2\#1^2 + k_3\#1^3, 1]$ denotes the minimum real root of the equation $[k_1x + k_2x^2 + k_3x^3 = 0$ in x and $DiracDelta[t]$ denotes Dirac delta function $\delta(t)$.

Appendix A: Proof of the Existence of Five Real Negative Roots

We shall prove that both $F(s)$ and $G(s)$ have five real negative roots. From Eq. (9):

$$F(s) = (s + l_1)(s + l_2)(s + l_3)F_1(s),$$

$$G(s) = (s + m_1)(s + m_2)(s + m_3)G_1(s),$$

where

$$F_1(s) = aa m_3(s + m_1)(s + m_2) + s(-a_2(m_2 - m_3)(s + m_1) - a_1(m_1 - m_3)(s + m_2)),$$

$$G_1(s) = bb l_3(s + l_1)(s + l_2) + s(-b_2(l_2 - l_3)(s + l_1) - b_1(l_1 - l_3)(s + l_2)).$$

In PET experiments, we can reasonably postulate $l_1 > l_2 > l_3 > 0$, and, $m_1 > m_2 > m_3 > 0$ because the radioactivity eventually approaches an equilibrium (the finite value). With respect to $F_1(s)$, we can see the following relationships:

$$F_1(0) = aa m_1 m_2 m_3,$$

$$F_1(-m_1) = a_1 m_1 (m_2 - m_1)(m_1 - m_3), \quad F_1(-m_3) = -(a_1 + a_2 + aa)m_3(m_1 - m_3)(m_3 - m_2).$$

As seen in Fig. 3, $aa > 0$ because the radioactivity is never negative even when $t \rightarrow \infty$. Furthermore, the largest and the smallest time constants: $1/m_3$ and $1/m_1$ correspond to the sampling data near the equilibrium and the initial stage, respectively, leading to the coefficient relations of the exponentials, $\exp(-m_3 t)$ and $\exp(-m_1 t)$: $-(a_1 + a_2 + aa) > 0$ and $a_1 < 0$, respectively. These facts lead to $F_1(0) > 0$, $F_1(-m_3) < 0$ and $F_1(-m_1) > 0$, showing that $F_1(s)$ has two real negative distinct roots, and then $F(s)$ has five real negative roots. Likewise, $G(s)$ has five real negative roots. \square

Appendix B: The Third-Order Polynomial in k_{b3}

In §2.4, we have derived the third-order polynomial by calculating the elimination ideal w.r.t. k_{b3} . This calculation needed 35.4 hours CPU time and 220 MBytes memory via Mathematica 5.2 (Wolfram Research, Inc.) with Intel(R) Xeon(R) CPU 2.33GHz. The calculated polynomial is as follows:

$$\begin{aligned}
 & (-r_3 + r_4)(r_3 - r_5)(r_4 - r_5)(r_1 k_4(r_4 - k_6)(r_5 - k_6)(k_{b3} + k_{p2}) - k_6 k_{b3} k_{p2} + k_6 k_{b3} k_{p2} + k_{p2} + 1) + (r_4(-k_{b3} k_{p2}) + r_5(k_{b3} + k_{p2} - k_{p6})) - (-k_{b3} k_{p2}) + r_5(k_{b3} + k_{p2}) r_5 - \\
 & k_{p6} - (r_4 - r_5) k_{b3} k_{p2} (r_5 - k_{p6}) + r_5 k_{b3} k_{p2} (r_5 - k_{b3}) (r_5 - k_{p2}) + r_4 r_5 (-k_{b3} k_{p2}) + r_5 (k_{b3} + k_{p2} - k_{p6}) - (-k_{b3} k_{p2}) + r_5 (k_{b3} + k_{p2}) (r_5 - k_{p6}) - (r_5 - r_5) k_{b3} \\
 & k_{p2} (r_5 - k_{p6}) - r_5^2 (r_4 (-k_{b3} k_{p2}) + r_5 (k_{b3} + k_{p2} - k_{p6})) + k_{b3} k_{p2} (-r_5 + k_{p6}) - (r_1 k_4 (r_5 - k_6) k_{b3} k_{p2} - k_6 k_{b3} k_{p2} - k_{p2}) + r_4 (r_5 (-k_{b3} k_{p2}) + r_5 (k_{b3} + k_{p2} \\
 & - k_{p6}) - (-k_{b3} k_{p2}) + r_5 (k_{b3} + k_{p2}) (r_5 - k_{p6})) - (r_5 - r_5) k_{b3} k_{p2} (r_5 - k_{p6}) + r_5^2 (r_4 (r_1 k_4 k_{b3} k_{p2} (r_4 k_{b3} k_{p2} - r_5 (k_{b3} + k_{p2} - k_{p6})) + k_{b3} k_{p2} (r_5 - k_{p6}) + r_5^2 \\
 & (-r_5^2 (r_5 - k_{b3}) (r_5 - k_{p2}) - (r_5 - r_1 - k_4 + k_6 - k_{b3}) k_{b3} k_{p2} (r_5 - k_{p6}) + r_4 (r_5^2 (k_{b3} + k_{p2}) + k_{b3} k_{p2} (r_1 + k_4 - k_6 + k_{b3} + k_{p6}) - r_5 (-k_6 k_{b3}) + k_{b3}^2 - k_6 k_{p2} + 2 k_{b3} \\
 & k_{p2} + r_1 (k_{b3} + k_{p2} - k_{p6}) + k_4 (k_{b3} + k_{p2} - k_{p6}) + k_6 k_{p6} + k_{p2} k_{p6})) + r_5 (r_5^2 k_{b3} (r_5 - k_{b3}) (r_5 - k_{p2}) - k_{b3} (-r_5 k_6) + k_4 k_{b3} + r_1 (k_4 + k_{b3}) k_{p2} (r_5 - k_{p6}) + r_4 (- \\
 & (r_5^2 k_6 (k_{b3} + k_{p2}) - k_{b3} k_{p2} (k_4 k_{b3} + r_1 (k_4 + k_{b3}) + k_6 k_{p6}) + r_5 (k_4 k_{b3} (k_{b3} + k_{p2} - k_{p6}) + r_1 (k_4 + k_{b3}) k_{b3} + k_{p2} - k_{p6}) + k_6 (k_{p2} k_{p6} + k_{b3} k_{p2} + k_{p6}))) \\
 & + r_4 (r_1 k_4 k_{b3} (-r_5^2 (r_5 - k_{b3}) (r_5 - k_{p2}) - (r_5 + k_6) k_{b3} k_{p2} (r_5 - k_{p6}) + r_4 (r_5^2 (k_{b3} + k_{p2}) + k_{b3} k_{p2} (-k_6 + k_{p6}) - r_5 (-k_6 (k_{b3} + k_{p2} - k_{p6})) + k_{p2} k_{p6} + k_{b3} \\
 & (k_{p2} + k_{p6}))) + r_5^2 (r_5^2 k_6 (r_5 - k_{b3}) (r_5 - k_{p2}) - k_{b3} (-r_5 k_6) + k_4 k_{b3} + r_1 (k_4 + k_{b3}) k_{p2} (r_5 - k_{p6}) + r_4 (-r_5^2 k_6 (k_{b3} + k_{p2}) - k_{b3} k_{p2} (k_4 k_{b3} + r_1 (k_4 + k_{b3}) + k_6 k_{p6}) \\
 & + r_5 (k_4 k_{b3} (k_{b3} + k_{p2} - k_{p6}) + r_1 (k_4 + k_{b3}) k_{b3} + k_{p2} - k_{p6}) + k_6 (k_{p2} k_{p6} + k_{b3} k_{p2} + k_{p6}))) + r_5 (r_5^2 (r_5 - k_{b3}) (-k_6 k_{b3}) + k_4 (-k_6 + k_{b3}) + r_1 (k_4 - k_6 + \\
 & k_{b3}) (r_5 - k_{p2}) + k_{b3} (k_4 k_6 k_{b3} + r_5 (-k_6 k_{b3}) + k_4 (-k_6 + k_{b3}) + r_1 (k_4 - k_6 + k_{b3})) + r_1 (k_6 k_{b3} + k_4 (k_6 + k_{b3})) k_{p2} (r_5 - k_{p6}) + r_4 (r_5^2 (k_4 (k_6 - k_{b3}) + k_6 k_{b3} - r_1 (k_4 - k_6 \\
 & + k_{b3}) k_{b3} + k_{p2}) + k_{b3} k_{p2} (k_4 k_6 k_{b3} + k_4 k_6 k_{p6} - k_4 k_{b3} k_{p6} + k_6 k_{b3} k_{p6} + r_1 (k_4 (k_6 + k_{b3} - k_{p6}) - k_{b3} k_{p6} + k_6 (k_{b3} + k_{p6}))) - r_5 (k_6 k_{b3} (k_{p2} k_{p6} + k_{b3} \\
 & (k_{p2} + k_{p6})) + k_4 (k_6 (k_{b3}^2 + 2 k_{b3} k_{p2} + k_{p2} k_{p6}) - k_{b3} (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6}))) + r_1 (k_4 (k_{b3}^2 + k_6 (k_{b3} + k_{p2} - k_{p6}) - 2 k_{b3} k_{p6} - k_{p2} k_{p6}) + k_6 (k_{b3}^2 + 2 \\
 & k_{b3} k_{p2} + k_{p2} k_{p6}) - k_{b3} (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6})))) - r_3 (r_1 k_4 (r_5 - k_6) k_{b3} k_{p2} (r_5^2 (r_5 - k_{b3}) (r_5 - k_{p2}) + r_4 (r_5 (-k_{b3} k_{p2}) + r_5 (k_{b3} + k_{p2} - k_{p6})) - (-k_{b3} \\
 & k_{p2}) + r_5 (k_{b3} + k_{p2}) (r_5 - k_{p6})) - (r_5 - r_5) k_{b3} k_{p2} (r_5 - k_{p6}) - (r_5^2 (r_4 (-k_{b3} k_{p2}) + r_5 (k_{b3} + k_{p2} - k_{p6})) + k_{b3} k_{p2} (-r_5 + k_{p6})) (r_1 k_4 k_{b3} (-r_5^2 (r_5 - k_{b3}) (r_5 \\
 & - k_{p2}) - (r_5 + k_6) k_{b3} k_{p2} (r_5 - k_{p6}) + r_4 (r_5^2 (k_{b3} + k_{p2}) + k_{b3} k_{p2} (-k_6 + k_{p6} \\
 & - r_5 (-k_6 (k_{b3} + k_{p2} - k_{p6})) + k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6}))) + r_5^2 (r_4 (-k_{b3} k_{p2}) + r_5 (k_{b3} + k_{p2} - k_{p6})) + k_{b3} k_{p2} (-r_5 + k_{p6})) (r_1 k_4 k_{b3} (-r_5^2 (r_5 - k_{b3}) (r_5 \\
 & - k_{p2}) - (r_5 + k_6) k_{b3} k_{p2} (r_5 - k_{p6}) + r_4 (r_5^2 (k_{b3} + k_{p2}) + k_{b3} k_{p2} (-k_6 + k_{p6} \\
 & - r_5 (-k_6 (k_{b3} + k_{p2} - k_{p6})) + k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6}))) + r_5 (r_5^2 (r_5 - k_{b3}) (-k_6 k_{b3}) + k_4 (-k_6 + k_{b3}) + r_1 (k_4 - k_6 + k_{b3})) (r_5 - k_{p2}) + k_{b3} (k_4 k_6 k_{b3} + r_5 \\
 & (-k_6 k_{b3}) + k_4 (-k_6 + k_{b3}) + r_1 (k_4 - k_6 + k_{b3})) + r_1 (k_6 k_{b3} + k_4 (k_6 + k_{b3})) k_{p2} (r_5 - k_{p6}) + r_4 (r_5^2 (k_4 (k_6 - k_{b3}) + k_6 k_{b3} - r_1 (k_4 - k_6 + k_{b3}) \\
 & k_{b3} + k_4 k_6 k_{p6} - k_4 k_{b3} k_{p6} + k_6 k_{b3} k_{p6} + r_1 (k_4 (k_6 + k_{b3} - k_{p6}) - k_{b3} k_{p6} + k_6 (k_{b3} + k_{p6}))) - r_5 (k_6 k_{b3} (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6})) + k_4 (k_6 (k_{b3}^2 + 2 k_{b3} \\
 & k_{p2} + k_{p2} k_{p6}) - k_{b3} (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6}))) + r_1 (k_4 (k_{b3}^2 + k_6 (k_{b3} + k_{p2} - k_{p6}) - 2 k_{b3} k_{p6} - k_{p2} k_{p6}) + k_6 (k_{b3}^2 + 2 k_{b3} k_{p2} + k_{p2} k_{p6}) - k_{b3} (k_{p2} \\
 & k_{p6} + k_{b3} (k_{p2} + k_{p6})))) + r_4 (-r_1 k_4 k_{b3} (r_5^2 (r_5 - k_{b3}) (r_5 - k_{p2}) + (r_5 + k_6) k_{b3} k_{p2} (r_5 - k_{p6}) + r_4 (-r_5^2 (k_{b3} + k_{p2}) + k_{b3} k_{p2} (k_6 - k_{p6}) + r_5 (-k_6 (k_{b3} + k_{p2} \\
 & - k_{p6})) + k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6}))) + r_5^2 (r_4 (-k_{b3} k_{p2}) + r_5 (k_{b3} + k_{p2} - k_{p6})) + k_{b3} k_{p2} (-r_5 + k_{p6})) (r_5^2 (r_5 - k_{b3}) (-k_6 k_{b3}) + k_4 (-k_6 + k_{b3}) + r_1 (k_4 - k_6 + \\
 & k_{b3}) (r_5 - k_{p2}) + k_{b3} (k_4 k_6 k_{b3} + r_5 (-k_6 k_{b3}) + k_4 (-k_6 + k_{b3}) + r_1 (k_4 - k_6 + k_{b3})) + r_1 (k_6 k_{b3} + k_4 (k_6 + k_{b3})) k_{p2} (r_5 - k_{p6}) + r_4 (r_5^2 (k_4 (k_6 - k_{b3}) + k_6 k_{b3} - r_1 (k_4 - k_6 \\
 & + k_{b3}) k_{b3} + k_{p2}) + k_{b3} k_{p2} (k_4 k_6 k_{b3} + k_4 k_6 k_{p6} - k_4 k_{b3} k_{p6} + k_6 k_{b3} k_{p6} + r_1 (k_4 (k_6 + k_{b3} - k_{p6}) - k_{b3} k_{p6} + k_6 (k_{b3} + k_{p6}))) - r_5 (k_6 k_{b3} (k_{p2} k_{p6} + k_{b3} \\
 & (k_{p2} + k_{p6})) + k_4 (k_6 (k_{b3}^2 + 2 k_{b3} k_{p2} + k_{p2} k_{p6}) - k_{b3} (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6}))) + r_5 (r_5^2 (r_5 - k_{b3}) (-k_4 + k_6) (k_6 - k_{b3}) + r_1 (k_4 - k_6 + k_{b3})) (r_5 - k_{p2})^2 + k_{b3}^2 (r_5^2 (-k_4 + k_6) (k_6 - k_{b3} \\
 & + r_1 (k_4 - k_6 + k_{b3})) + k_6 (k_4 k_6 k_{b3} + r_1 (k_4 k_6 + 2 k_4 k_{b3} + k_6 k_{b3})) + r_5 (-k_6 (k_4 (k_6 - k_{b3}) + k_6 k_{b3})) + r_1 (k_6 (-k_6 + k_{b3}) + k_4 (k_6 + 2 k_{b3}))) k_{p2}^2 (r_5 - k_{p6})^2 - r_5^2 (r_5 - k_{b3} \\
 &) (r_5 - k_{p2}) + 2 r_5^2 (-k_4 + k_6) (k_6 - k_{b3}) + r_1 (k_4 - k_6 + k_{b3}) (k_{b3} + k_{p2}) + k_{b3} k_{p2} (r_1 ((k_6 - k_{b3}) k_6 - 2 k_{p6}) - k_4 (k_6 + 2 k_{b3} - 2 k_{p6})) + k_4 (k_6 - k_{b3}) (k_6 - 2 k_{p6}) + k_6 (k_6 k_{b3} \\
 & + 2 k_6 k_{p6} - 2 k_{b3} k_{p6})) + r_5 (-k_4 (k_6 - k_{b3}) (k_{b3} + k_{p2} - k_{p6}) - 2 (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6}))) - k_6 (-2 k_{b3} (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6})) + k_6 (k_{b3}^2 + 2 k_{p2} k_{p6} \\
 & + k_{b3} (3 k_{p2} + k_{p6}))) + r_1 (k_4 (k_6 (k_{b3} + k_{p2} - k_{p6}) + 2 (k_{b3}^2 - 2 k_{b3} k_{p6} - k_{p2} k_{p6})) - (k_6 - k_{b3}) (k_6 (k_{b3} + k_{p2} - k_{p6}) - 2 (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6})))) + r_4 k_{b3} \\
 & k_{p2} (r_5 - k_{p6}) - 2 r_5^2 (-k_4 + k_6) (k_6 - k_{b3}) + r_1 (k_4 - k_6 + k_{b3}) (k_{b3} + k_{p2}) + k_{b3} k_{p2} (k_6 (k_6 k_{b3} k_{p6} + k_4 (2 k_6 k_{b3} + k_6 k_{p6} - k_{b3} k_{p6})) + r_1 (k_4 (k_6 + 2 k_{b3}) (2 k_6 - k_{p6} \\
 &) + k_6 (2 k_6 k_{b3} + k_6 k_{p6} - k_{b3} k_{p6})) + r_5 (-k_6 k_{b3} (-2 k_{b3} k_{p2} k_{p6} + k_6 (3 k_{p2} k_{p6} + k_{b3} (2 k_{p2} + k_{p6}))) + k_4 (-2 k_{b3}^2 k_{p2} k_{p6} - k_6^2 (2 k_{b3}^2 + 4 k_{b3} k_{p2} - k_{b3} k_{p6} \\
 & + k_{p2} k_{p6}) + k_6 k_{b3} (3 k_{p2} k_{p6} + k_{b3} (2 k_{p2} + k_{p6}))) + r_1 (-2 k_{b3}^2 k_{p2} k_{p6} - k_6^2 (2 k_{b3}^2 + 4 k_{b3} k_{p2} - k_{b3} k_{p6} + k_{p2} k_{p6}) + k_6 k_{b3} (2 k_{b3} k_{p2} + k_{b3} k_{p6} + 3 \\
 & k_{p2} k_{p6}) + k_4 (-2 k_6^2 (k_{b3} + k_{p2} - k_{p6}) + 2 k_{b3}^2 (2 k_{p2} + k_{p6}) + k_6 (-4 k_{b3}^2 - 2 k_{b3} k_{p2} + 5 k_{b3} k_{p6} + k_{p2} k_{p6}))) - r_5^2 (-k_4 (k_6 - k_{b3}) (k_6 (2 k_{b3} + 2 k_{p2} - k_{p6}) - 2 (k_{p2} \\
 & k_{p6} + k_{b3} (k_{p2} + k_{p6}))) - k_6 (k_6 (k_{b3} + 2 k_{p2}) (2 k_{b3} + k_{p6}) - 2 k_{b3} (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6}))) + r_1 (k_4 (2 (2 k_{b3} + k_{p2}) (k_{b3} - k_{p6}) + k_6 (2 k_{b3} + 2 k_{p2} - k_{p6})) - (k_6 \\
 & - k_{b3}) (k_6 (2 k_{b3} + 2 k_{p2} - k_{p6}) - 2 (k_{p2} k_{p6} + k_{b3} (k_{p2} + k_{p6})))) + r_4^2 (r_5^2 (-k_4 + k_6) (k_6 - k_{b3}) + r_1 (k_4 - k_6 + k_{b3})) (k_{b3}^2 + 4 k_{b3} k_{p2} + k_{p2}^2) + k_{b3}^2 k_{p2}^2 (k_6 k_{p6}
 \end{aligned}$$

$$\begin{aligned}
 & (-k_{h3}k_{p0}) + k_6(2k_{h3} + k_{p0}) + k_4(k_{h3}k_{p0}^2 - k_0k_{p0}(2k_{h3} + k_{p0}) + k_0^2(k_{h3} + 2k_{p0})) + t_1(k_{h3}(k_{p0}^2 - k_0k_{p0}(2k_{h3} + k_{p0})) + k_0^2(k_{h3} + 2k_{p0})) + k_4(k_0^2 + 2k_0 \\
 & (k_{h3} - k_{p0}) + k_{p0}(-4k_{h3} + k_{p0}))) - t_5^2k_{h3}k_{p2}(k_4(2k_{h3}k_{p0}(k_{p2}k_{p0} + k_{h3}(2k_{p2} + k_{p0})) + k_0^2(2k_{h3}^2 + (3k_{p2} - k_{p0})k_{p0} + k_{h3}(4k_{p2} + k_{p0})) - k_0(2k_{p2} \\
 & k_{p0}^2 + k_{h3}k_{p0}(7k_{p2} + k_{p0}) + k_{p0}^2(2k_{p2} + 3k_{p0}))) + k_0(-2k_{h3}k_{p0}(k_{p2}k_{p0} + k_{h3}(2k_{p2} + k_{p0})) + k_0(2k_{p2}k_{p0}^2 + k_{h3}k_{p0}(7k_{p2} + k_{p0})) + k_{p0}^2(2k_{p2} + 3k_{p0} \\
 &))) + t_1(2k_{h3}k_{p0}(k_{p2}k_{p0} + k_{h3}(2k_{p2} + k_{p0})) + k_0^2(2k_{h3}^2 + (3k_{p2} - k_{p0})k_{p0} + k_{h3}(4k_{p2} + k_{p0})) - k_0(2k_{p2}k_{p0}^2 + k_{h3}k_{p0}(7k_{p2} + k_{p0})) + k_{p0}^2(2k_{p2} + 3 \\
 & k_{p0})) + k_4(2k_0^2(k_{h3} + k_{p2} - k_{p0}) - 2k_{h3}(k_{p2} - 2k_{p0})k_{p0} + 2k_{p2}k_{p0}^2 - 2k_{h3}^2(2k_{p2} + 3k_{p0})) + k_6(4k_{h3}^2 + 2k_{h3}k_{p2} - 7k_{h3}k_{p0} - 3k_{p2}k_{p0} + k_{p0}^2(1))) + t_5^2(-t_1 \\
 & k_4(k_0 - k_{h3})k_{h3}(k_{h3}^2 + 3k_{h3}k_{p2} + k_{p2}^2 - k_{h3}k_{p0} - k_{p2}k_{p0}) - 2(k_{p2}^2k_{p0} + k_{h3}^2(2k_{p2} + k_{p0}) + k_{h3}k_{p2}(2k_{p2} + 3k_{p0}))) - k_0(k_0(k_{h3}^3 + 2k_{p2}^2k_{p0} + 5k_{h3} \\
 & k_{p2}(k_{p2} + k_{p0}) + k_{h3}^2(7k_{p2} + k_{p0})) - 2k_{h3}(k_{p2}^2k_{p0} + k_{h3}^2(2k_{p2} + k_{p0}) + k_{h3}k_{p2}(2k_{p2} + 3k_{p0}))) + t_1((-k_0 + k_{h3})k_{h3}(k_{h3}^2 + 3k_{h3}k_{p2} + k_{p2}^2 - k_{h3} \\
 & k_{p0} - k_{p2}k_{p0}) - 2(k_{p2}^2k_{p0} + k_{h3}^2(2k_{p2} + k_{p0}) + k_{h3}k_{p2}(2k_{p2} + 3k_{p0}))) + k_4(k_0(k_{h3}^2 + 3k_{h3}k_{p2} + k_{p2}^2 - k_{h3}k_{p0} - k_{p2}k_{p0}) + 2(k_{h3}^3 + k_{h3}^2(k_{p2} - 2 \\
 & k_{p0}) - k_{p2}^2k_{p0} - k_{h3}k_{p2}(k_{p2} + 4k_{p0}))) + t_5^2(k_0(k_{p2}^2k_{p0} + k_{h3}^2(3k_{p2} + k_{p0}) + k_{h3}k_{p2}(7k_{p2} + k_{p0})) + 2k_{h3}^2k_{p2}(3k_{p2} + 4k_{p0})) - k_{h3}(k_{p2}^2 \\
 & k_{p0}^2 + 2k_{h3}k_{p2}k_{p0}(3k_{p2} + k_{p0}) + k_{h3}^2(3k_{p2}^2 + 6k_{p2}k_{p0} + k_{p0}^2))) + k_4(k_0^2(k_{h3}^3 + 4k_{h3}k_{p2}^2 + k_{h3}^2(5k_{p2} - k_{p0}) + k_{p2}(k_{p2} - k_{p0})k_{p0}) - k_0(k_{p2}^2 \\
 & k_{p0}^2 + k_{h3}^3(3k_{p2} + k_{p0}) + k_{h3}k_{p2}(7k_{p2} + k_{p0}) + 2k_{h3}^2k_{p2}(3k_{p2} + 4k_{p0})) + k_{h3}(k_{p2}^2k_{p0}^2 + 2k_{h3}k_{p2}k_{p0}(3k_{p2} + k_{p0}) + k_{h3}^2(3k_{p2}^2 + 6k_{p2} \\
 & k_{p0} + k_{p0}^2))) + t_1(k_0^2(k_{h3}^3 + 4k_{h3}k_{p2}^2 + k_{h3}^2(5k_{p2} - k_{p0}) + k_{p2}(k_{p2} - k_{p0})k_{p0}) - k_0(k_{p2}^2k_{p0}^2 + k_{h3}^3(3k_{p2} + k_{p0}) + k_{h3}k_{p2}(7k_{p2} + k_{p0})) + \\
 & 2k_{h3}^2k_{p2}(3k_{p2} + 4k_{p0})) + k_{h3}(k_{p2}^2k_{p0}^2 + 2k_{h3}k_{p2}k_{p0}(3k_{p2} + k_{p0}) + k_{h3}^2(3k_{p2}^2 + 6k_{p2}k_{p0} + k_{p0}^2))) + k_4(k_0^2(k_{h3} + k_{p2} - k_{p0})^2 + k_{p2}^2k_{p0}^2 + 4 \\
 & k_{h3}k_{p2}k_{p0}(k_{p2} + k_{p0}) - 2k_{h3}^2(3k_{p2} + k_{p0}) + k_{h3}^2(-3k_{p2}^2 + 2k_{p2}k_{p0} + 3k_{p0}^2) + k_6(2k_{h3}^3 + k_{h3}^2(k_{p2} - 5k_{p0}) + k_{p2}k_{p0}(-k_{p2} + k_{p0}) - k_{h3}(k_{p2}^2 + 6 \\
 & k_{p2}k_{p0} - 3k_{p0}^2))))))K_1/K_{p1}.
 \end{aligned}$$

Victor G. Ganzha Ernst W. Mayr
Evgenii V. Vorozhtsov (Eds.)

Computer Algebra in Scientific Computing

10th International Workshop, CASC 2007
Bonn, Germany, September 16-20, 2007
Proceedings

 Springer

5. Goldwasser, S.: The Search for Provably Secure Cryptosystems. In: Pomerance, C. (ed.) *Cryptology and Computational Number Theory. Proceedings of Symposia in Applied Mathematics*, vol. 42, American Mathematical Society (1990)
6. Goldwasser, S., Micali, S.: Probabilistic Encryption. *Journal of Computer and System Sciences* 28, 270-299 (1984)
7. McCurley, K.S.: Odds and Ends from Cryptology and Computational Number Theory. In: Pomerance, C. (ed.) *Cryptology and Computational Number Theory. Proceedings of Symposia in Applied Mathematics*, vol. 42, pp. 49-74. American Mathematics Society (1990)
8. Yan, S.Y.: A New Cryptographic Scheme based on the k th Power Residuosity Problem. In: 15th British Colloquium for Theoretical Computer Science (BCTCS15), 14-16 April 1999, Keele University (1999)
9. Yau, S.Y.: *Number Theory for Computing*, 2nd edn. Springer, Heidelberg (2002)

An Algebraic-Numeric Algorithm for the Model Selection in Kinetic Networks

Hiroshi Yoshida^{1,*}, Koji Nakagawa², Hirokazu Anai³, and Katsuhisa Horimoto^{2,*}

¹ Faculty of Mathematics, Organization for the Promotion of Advanced Research, Kyushu University, Hakozaki 6-10-1, Higashi-ku, Fukuoka 812-8581 Japan
Tel/Fax.: +81-92-642-7396
hiroshi@math.kyushu-u.ac.jp

² Computational Biology Research Centre (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), Aomi 2-42, Koto-ku, Tokyo 135-0064, Japan
* Tel.: +81-3-3599-8711; Fax.: +81-3-3599-8081
{nakagawa-koji, k.horimoto}@aist.go.jp

³ IT Core Laboratories, FUJITSU LABORATORIES LTD./CREST, JST., Kamikodanaka 4-1-1, Nakahara-ku, Kawasaki 211-8588, Japan
anai@jp.fujitsu.com

Abstract. We propose a novel algorithm to select a model that is consistent with the time series of observed data. In the first step, the kinetics for describing a biological phenomenon is expressed by a system of differential equations, assuming that the relationships between the variables are linear. Simultaneously, the time series of the data are numerically fitted as a series of exponentials. In the next step, both the system of differential equations with the kinetic parameters and the series of exponentials fitted to the observed data are transformed into the corresponding system of algebraic equations, by the Laplace transformation. Finally, the two systems of algebraic equations are compared by an algebraic approach. The present method estimates the model's consistency with the observed data and the determined kinetic parameters. One of the merits of the present method is that it allows a kinetic model with cyclic relationships between variables that cannot be handled by the usual approaches. The plausibility of the present method is illustrated by the actual relationships between specific leaf area, leaf nitrogen and leaf gas exchange with the corresponding simulated data.

1 Introduction

The knowledge-based approach to constructing a biological network model is recognized as one of the most promising approaches [4]. In this approach, the causal relations between biological molecules are described as a directed graph, based on the gene interaction information collected from a large number of previous reports. Since each relation identified by experimental studies is regarded as strong evidence for the existence of edges in the network model, biological network models have been constructed for various biological phenomena by a knowledge-based approach. On the other hand, it is well-known that the relationships between the molecules in a living cell change

* Corresponding authors.

dynamically, depending on the cellular environment. Thus, the molecular relationships in the literature represent the responses to the different conditions in the experimental studies, and in the network model generated from the biological knowledge, the consistency of the model with the data observed by experimental studies must be considered carefully. Actually, several distinctive models of the relationship between molecules for a biological phenomenon can be obtained from the large amount of information in the literature [2, 5]. In these cases, a model that is consistent with the data observed under particular conditions should be selected from the candidate models.

The consistency of a model with the observed data first reminds us of the identifiability problem in the compartmental models for tracer kinetics [1, 5, 6]. In the compartmental models, the unknown parameters are estimated from tracer data in the accessible pools. The identifiability problem addresses the issue of whether the unknown parameters can be determined uniquely or non-uniquely from the tracer data. This issue has usually been solved through the transformation of differential equations into algebraic equations, by the Laplace transformation. Although a systematic algorithm for the identifiability problem was proposed [3], its application is limited to the unrealistic context of an error-free model structure and noise-free tracer data. Thus, it still seems to be difficult to solve the identifiability problem for actually observed data, in spite of the mathematical studies.

The issue of the consistency of a model with the observed data is also well known in statistics, as the test for causal hypotheses by using the observed data. The origin of the test for causal hypotheses is attributed to path analysis [12]. Unfortunately, the importance of this cornerstone research has been ignored for a long time, but the natural extension of the path analysis has been established as the well-known structural equation model (SEM) [8]. Indeed, the SEM has been utilized recently in various fields, in accordance with increased computer performance. However, the SEM without any latent variables, which is the natural form for applying the SEM to the biological networks, frequently faces difficulty in the numerical calculation of the maximum likelihood for the observed data. To overcome the difficulty of this calculation, the d-sep test [11] has been developed, based on the concept of d-separation in a directed acyclic graph [10]. Notice that the graph consistency with the data in the d-sep test can consider only the directed acyclic graph (DAG), without any cyclic relationships.

In this study, we propose a new method for selecting models, by estimating the consistency of a kinetic model with the time series of observed data. Our method is described in the following section. First, the kinetics for describing a biological phenomenon is expressed by a system of differential equations, assumed that the relationships between the variables are linear. Simultaneously, the time series of the data are numerically fitted as a series of exponentials. Next, the differential equations with the kinetic parameters and the series of exponentials fitted to the observed data are both transformed into the corresponding system of algebraic equations, by the Laplace transformation. Finally, the two systems of algebraic equations are compared by an algebraic approach. Thus, the present method estimates the model's consistency with the observed data and the determined kinetic parameters. In §3, the plausibility of the present method is illustrated by the actual relationships between specific leaf area, leaf

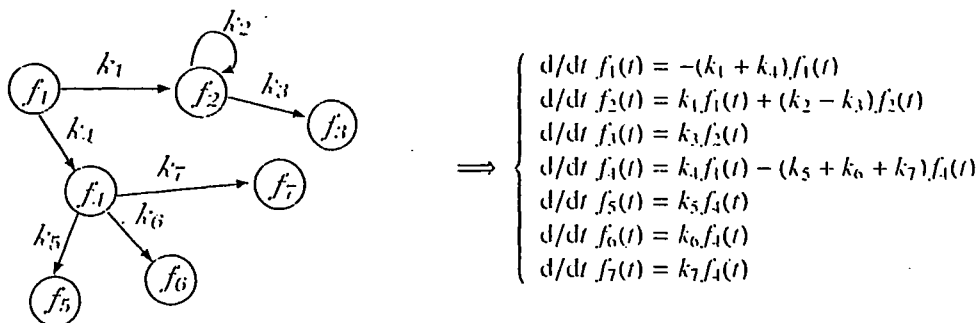


Fig. 1. Correspondence between a network and a system of differential equations. By assuming a linear relation between the variables, the kinetics of chemicals f_1, f_2, \dots in the left graph can be described by the system of differential equations on the right side.

nitrogen and leaf gas exchange [9], with the corresponding data generated by the differential equations for the relationships. Furthermore, the merits and pitfalls of the present method are discussed. In particular, one of the merits of the present method is that it allows a kinetic model with cyclic relationships between variables that cannot be handled by the usual approaches.

2 Methods

The aim of this paper is to select the model most consistent with the given sampling data. In this section, we propose a method to perform this selection, where the model is described as a network. The network addressed in this paper designates the kinetics of chemicals, which can be described by a system of differential equations, as seen in Fig. 1.

First, we will show the overview of our method by a schematic illustration. We will then provide an explanation for the Laplace transformations of model formulae and sampling data over the time domain, as preparation for the model selection over the Laplace domain. Lastly, we describe a procedure to estimate the model consistency with the definition of *consistency measure*.

2.1 Overview

The overview of our method is schematically illustrated in Fig. 2. The point is that we perform the model selection over the Laplace domain. Therefore, both the model formulae and sampling data must be transformed into functions over the Laplace domain. Suppose that the model formulae are $\{d/dt h(t) = -k_1 h(t), d/dt f(t) = k_1 h(t) - k_2 f(t)\}$ and the sampling data are fitted to $h_o(t) = \beta_0 \exp(-\alpha_0 t), f_o(t) = \beta_1 \exp(-\alpha_1 t) + \beta_2 \exp(-\alpha_2 t)$. The Laplace-transformed formulae of the model formula: $L[f(t)](s)$ and the fitted function: $L[f_o(t)](s)$ are rational functions in s , as seen in the middle row of Fig. 2. Let *comp* denote the set of polynomials obtained by matching the coefficients in s of $L[f(t)](s)$ and $L[f_o(t)](s)$ over the Laplace domain, in which every element is equal to zero when $L[f(t)](s)$ is exactly identical to $L[f_o(t)](s)$ in s . Then we have adopted the smallest sum-square value of the elements in *comp* as a *consistency measure* between the model

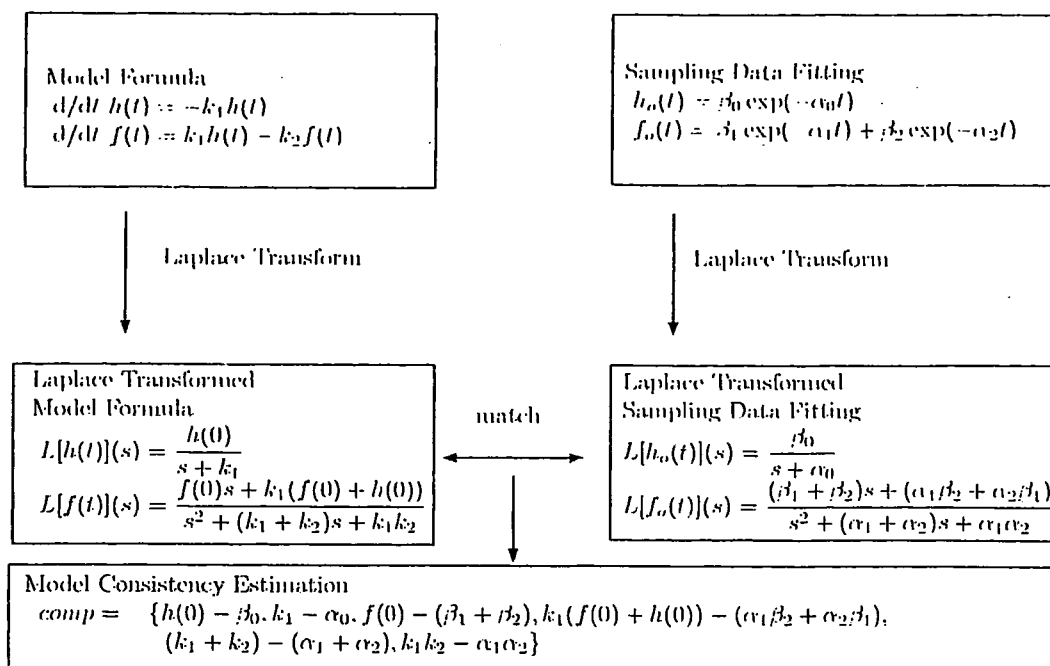


Fig. 2. Overview of our method. The top row designates the model formulae and the sampling data over the time domain, and the middle row designates their Laplace transformations. $comp$ denotes the set of polynomials derived by matching the coefficients in s of $L[f(t)](s)$ and $L[f_o(t)](s)$ over the Laplace domain, which is zero when the model and sampling data are completely consistent with each other.

and the sampling data, because this value is zero in the case of $L[f(t)](s) = L[f_o(t)](s)$. We shall mention the formal procedure and definitions concretely in the following subsections.

2.2 Preparations: Transformation into Laplace Domain

Model Formula. Suppose that the model formulae are described over the time domain as the following system of differential equations:

$$\frac{df_i(t)}{dt} = F_i(\vec{f}, \vec{k}), \quad (2.1)$$

where $\vec{f} = \{f_1, f_2, \dots, f_n\}$ and $\vec{k} = \{k_1, k_2, \dots, k_m\}$. $F_i(\vec{f}, \vec{k})$ can be determined in accordance with the network representing the model, and \vec{k} denotes the kinetic constants between the chemicals. We transform this system of differential equations into the system of algebraic equations over the Laplace domain, and solve the equations in $L[f_i(t)](s)$ ($i = 1, 2, \dots, n$). Notice that in this paper, we deal only with an autonomous system of differential equations, but in the framework of the Laplace transformation, we can deal with differential equations containing external forces or 'convolutions' of complex functions, as long as the Laplace-transformed algebraic equations can explicitly be solved in $L[f_i(t)](s)$ ($i = 1, 2, \dots, n$).

Sampling Data Fitting. In this paper, we need the Laplace transformation of the sampling data, because we perform the model selection over the Laplace domain. Let $f_{o_i}(t)$ denote the sampling data corresponding to $f_i(t)$ derived theoretically. By using non-linear regression (via Maple 10 Global Optimization toolbox, ©MapleSoft), $f_{o_i}(t)$ is expressed in terms of a series of exponentials, according to [6], as follows:

$$f_{o_i}(t) = \beta_0 + \sum_{i=1}^k \beta_i \exp(-\alpha_i t), \tag{2.2}$$

where k is the number of distinct exponentials determined by $f_i(t)$, and β_0 is zero in the case of the non-existence of a constant term within $f_i(t)$. $f_{o_i}(t)$ thus fitted is changed into the Laplace-transformed data as follows:

$$L[f_{o_i}(t)](s) = \frac{\beta_0}{s} + \sum_{i=1}^k \frac{\beta_i}{s + \alpha_i}, \tag{2.3}$$

where L denotes the Laplace transformation.

2.3 Estimation of Model Consistency

Consistency Measure. To evaluate the consistency of the model with the sampling data, here we define two *consistency measures*. If the model is completely consistent with the sampling data and the data lack noise and inaccuracies, then $L[f_i(t)](s) = L[f_{o_i}(t)](s)$ ($i = 1, 2, \dots, n$) holds. This fact has led us to the following definitions of consistency measure:

Let *comp* denote the set of polynomials obtained by matching the coefficients of $L[f(t)](s)$ and $L[f_{o}(t)](s)$ over the Laplace domain, in which every element is zero in the case of $L[f_i(t)](s) = L[f_{o_i}(t)](s)$ ($i = 1, 2, \dots, n$); that is, when Formula $L[f_i(t)](s) = L[f_{o_i}(t)](s)$ is an identity in s .

The first consistency measure (in short, *CM1*) of the model is defined as the smallest sum-square value of the elements in *comp* under the following constraint:

$$k_1 > 0, k_2 > 0, \dots, k_m > 0. \tag{2.4}$$

In order to obtain the smallest value, we have utilized the least squares method using the following equations:

$$\frac{\partial}{\partial k_1} g(\vec{k}) = 0, \frac{\partial}{\partial k_2} g(\vec{k}) = 0, \dots, \frac{\partial}{\partial k_m} g(\vec{k}) = 0, \tag{2.5}$$

where $g(\vec{k})$ is the sum-square value of the elements in *comp*. It should be noted that in this paper we deal only with the case that the ideal associated with the set of polynomials in (2.5) is zero-dimensional. Then, we survey all of the possible candidates of the minimum by calculating *all* of the real positive roots of the system of algebraic equations (2.5). Several methods and tools exist to calculate all real roots of algebraic equations adjoined by a zero-dimensional ideal. Here we employed ‘NSolve’ in Mathematica 5.2 (Wolfram Research Inc.), which computes the desired roots efficiently.

Remark 1. If $comp$ is a positive dimensional, then we can always perturb the set of polynomials in $comp$ in order to obtain a zero-dimensional variety. Although here we cannot discuss the stability and convergency issues related to such perturbations, it is an important research issue on its own light (see [7] for an example).

In this paper, we have calculated the other consistency measure (in short, $CM2$) as the smallest $g(\vec{k})$ under the following constraint:

$$k_1 \geq 0, k_2 \geq 0, \dots, k_m \geq 0. \quad (2.6)$$

The difference between Constraints: (2.4) and (2.6) is that one takes account of the zero value of the kinetic constants \vec{k} , corresponding to the non-existence of edges in the network. This account yields a finer model selection where all of the subnetworks of the presupposed network are also considered. We can calculate the smallest value of $g(\vec{k})$ under Constraint (2.6), using the following recursive definition:

Let $MinimumValue(q(\vec{l}))$ denote the *minimum value* of function q with variables: $\vec{l} = \{l_1, l_2, \dots, l_m\}$ by the following procedure:

1. If the cardinality of \vec{l} , namely m , is zero, then the *minimum value* is infinity.
2. Otherwise, let v_0 denote the minimum value of q under Constraint (2.4) via 'NSolve.' Furthermore, let v_i ($i = 1, 2, \dots, m$) denote the value calculated by $MinimumValue(q(\vec{l}_i))$, where \vec{l}_i is the vector: $\{l_1, l_2, \dots, l_{i-1}, 0, l_{i+1}, \dots, l_m\}$.
3. The *minimum value* is the smallest value among v_0, v_1, \dots, v_m .

Model Selection. Using the consistency measure defined in §2.3, we performed a model selection. We, first, calculated the consistency measures among all of the combinations of the presupposed models with the sampling data. Next, we arranged the combinations of the models with the data in ascending order by the consistency measure. Last, we estimated the most consistent model having the first element (the smallest value).

3 Results and Discussion

3.1 Preparations: Transformation into Laplace Domain

Model Formula. We analyzed the models for a relationship between specific leaf area, leaf nitrogen, and leaf gas exchange in botany [9]. In the original paper, six models for the kinetics of four biomolecules are listed, and the consistency of the models with the observed data, which are composed of various properties of the molecules, rather than time series data, are tested by the d-sep test. In this paper, four of the six original models (models A, B, C, and D) and one model (model E) modified from the original one are considered, to show how cyclic relationships can be handled. The models considered in this paper are shown in Fig. 3. Each model expressed the relationship between four biomolecules, SLA , N , A , and G . According to the definition in §2.2, each relationship

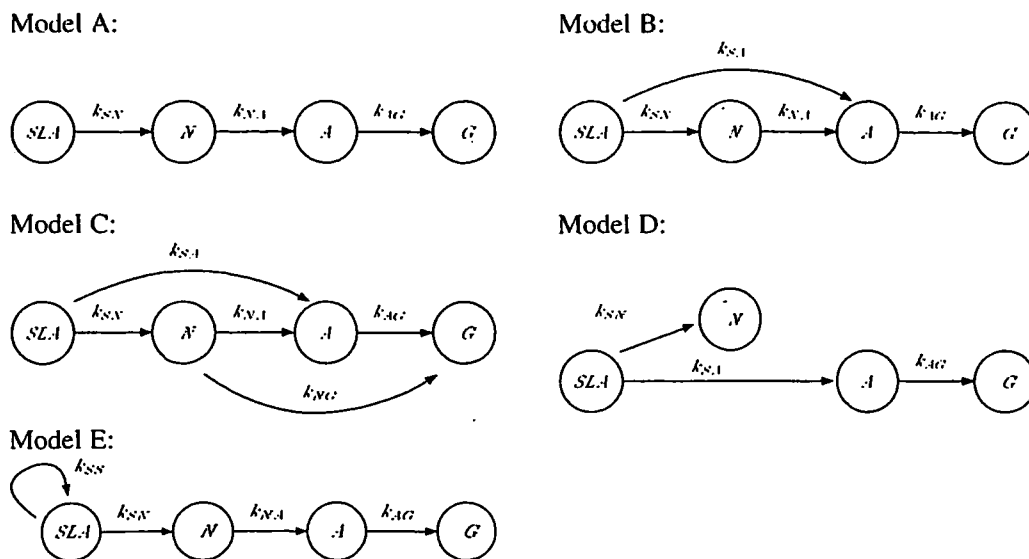


Fig. 3. Models analyzed in the present study. In the above models, the causal relationships between molecules are denoted by arrows. The molecules corresponding to the variables, denoted within the circles, are SLA , N , A , and G , and the kinetic parameters, denoted over the arrows, are k_{SN} , k_{NA} , k_{AG} , k_{SA} , k_{NG} , and k_{SS} .

between the variables is assumed to be linear, and then the differential equations for the five models can be formulated as follows:

Model A:

$$\begin{cases} d/dt SLA(t) = -k_{SN} SLA(t), \\ d/dt N(t) = k_{SN} SLA(t) - k_{NA} N(t), \\ d/dt A(t) = k_{NA} N(t) - k_{AG} A(t), \\ d/dt G(t) = k_{AG} A(t). \end{cases} \quad (3.1)$$

Model B:

$$\begin{cases} d/dt SLA(t) = -(k_{SN} + k_{SA}) SLA(t), \\ d/dt N(t) = k_{SN} SLA(t) - k_{NA} N(t), \\ d/dt A(t) = k_{SA} SLA(t) + k_{NA} N(t) - k_{AG} A(t), \\ d/dt G(t) = k_{AG} A(t). \end{cases} \quad (3.2)$$

Model C:

$$\begin{cases} d/dt SLA(t) = -(k_{SN} + k_{SA}) SLA(t), \\ d/dt N(t) = k_{SN} SLA(t) - (k_{NA} + k_{NG}) N(t), \\ d/dt A(t) = k_{SA} SLA(t) + k_{NA} N(t) - k_{AG} A(t), \\ d/dt G(t) = k_{AG} A(t) + k_{NG} N(t). \end{cases} \quad (3.3)$$

Model D:

$$\begin{cases} d/dt SLA(t) = -(k_{SN} + k_{SA}) SLA(t), \\ d/dt N(t) = k_{SN} SLA(t), \\ d/dt A(t) = k_{SA} SLA(t) - k_{AG} A(t), \\ d/dt G(t) = k_{AG} A(t). \end{cases} \quad (3.4)$$

Model E:

$$\begin{cases} d/dt SLA(t) = (k_{SS} - k_{SN}) SLA(t), \\ d/dt N(t) = k_{SN} SLA(t) - k_{NA} N(t), \\ d/dt A(t) = k_{NA} N(t) - k_{AG} A(t), \\ d/dt G(t) = k_{AG} A(t). \end{cases} \quad (3.5)$$

In the above equations, k_{SN} , k_{NA} , k_{AG} , k_{SA} , k_{NG} , and k_{SS} are the kinetic parameters between the molecules. Notice that the relationships between the molecules in the actual kinetics cannot be expressed by the above equations. In the actual case, some relationships are non-linear, such as the well-known Michaelis–Menten kinetics in enzyme reactions. In the present study, we have adopted the relationships between molecules as typical ones, but do not consider the details of the kinetics between molecules.

According to the definitions in §2.2, we transform the above systems of differential equations of (3.1)–(3.5) into the system of algebraic equations over the Laplace domain, and solve the equations for the five models. For instance, the solution to the system of differential equations for Model A is expressed over the Laplace domain, as follows:

$$\left\{ \begin{aligned} L[SLA(t)](s) &= \frac{SLA(0)}{s + k_{SN}}, \\ L[N(t)](s) &= \frac{N(0)s + N(0)k_{SN} + k_{SN}SLA(0)}{s^2 + (k_{SN} + k_{NA})s + k_{NA}k_{SN}}, \\ L[A(t)](s) &= \frac{A(0)s^2 + (k_{NA}N(0) + A(0)k_{SN} + A(0)k_{NA})s + k_{NA}N(0)k_{SN} + k_{NA}k_{SN}SLA(0) + A(0)k_{NA}k_{SN}}{(s^3 + (k_{AG} + k_{SN} + k_{NA})s^2 + (k_{AG}k_{SN} + k_{AG}k_{NA} + k_{NA}k_{SN})s + k_{AG}k_{NA}k_{SN})}, \\ L[G(t)](s) &= \frac{G(0)s^3 + (G(0)k_{AG} + G(0)k_{NA} + k_{AG}A(0) + G(0)k_{SN})s^2 + (G(0)k_{AG}k_{NA} + k_{AG}k_{NA}N(0) + G(0)k_{AG}k_{SN} + k_{AG}A(0)k_{NA} + k_{AG}A(0)k_{SN} + G(0)k_{NA}k_{SN})s + k_{AG}k_{NA}N(0)k_{SN} + k_{AG}k_{NA}k_{SN}SLA(0) + G(0)k_{AG}k_{NA}k_{SN} + k_{AG}A(0)k_{NA}k_{SN}}{(s^4 + (k_{AG} + k_{SN} + k_{NA})s^3 + (k_{AG}k_{SN} + k_{AG}k_{NA} + k_{NA}k_{SN})s^2 + s k_{AG}k_{NA}k_{SN})}. \end{aligned} \right. \quad (3.6)$$

In the above equations, the initial values for each molecule are denoted by $SLA(0)$, $N(0)$, $A(0)$, and $G(0)$.

Sampling Data Fitting. To estimate the consistency of the above equations derived from the models with the data, we should presuppose the equations for the sampling data. For this purpose, first, a series of exponentials with parameters are set. For instance, the equations for fitting to the data in Model A are expressed as follows:

$$\left\{ \begin{aligned} SLA_0(t) &= \beta_{SLA,1} \exp(-\alpha_{SLA,1}t), \\ N_0(t) &= \beta_{N,1} \exp(-\alpha_{N,1}t) + \beta_{N,2} \exp(-\alpha_{N,2}t), \\ A_0(t) &= \beta_{A,1} \exp(-\alpha_{A,1}t) + \beta_{A,2} \exp(-\alpha_{A,2}t) + \beta_{A,3} \exp(-\alpha_{A,3}t), \\ G_0(t) &= \beta_{G,1} \exp(-\alpha_{G,1}t) + \beta_{G,2} \exp(-\alpha_{G,2}t) + \beta_{G,3} \exp(-\alpha_{G,3}t) + \beta_{G,4}. \end{aligned} \right. \quad (3.7)$$

Then, the corresponding algebraic equations are obtained by the Laplace transformation. The corresponding algebraic equations in Model A are as follows:

$$\left\{ \begin{aligned} L[SLA_0(t)](s) &= \frac{\beta_{SLA,1}}{s + \alpha_{SLA,1}}, \\ L[N_0(t)](s) &= \frac{\beta_{N,1}}{s + \alpha_{N,1}} + \frac{\beta_{N,2}}{s + \alpha_{N,2}}, \\ L[A_0(t)](s) &= \frac{\beta_{A,1}}{s + \alpha_{A,1}} + \frac{\beta_{A,2}}{s + \alpha_{A,2}} + \frac{\beta_{A,3}}{s + \alpha_{A,3}}, \\ L[G_0(t)](s) &= \frac{\beta_{G,1}}{s + \alpha_{G,1}} + \frac{\beta_{G,2}}{s + \alpha_{G,2}} + \frac{\beta_{G,3}}{s + \alpha_{G,3}} + \frac{\beta_{G,4}}{s}. \end{aligned} \right. \quad (3.8)$$

Notice that the parameters in the above equations are estimated by numerically fitting them to the data.

3.2 Estimation of Model Consistency

Data Generation for Simulation. In the present study, we have no actual data for the molecules in the models, and thus we need to generate the time series of data for the constituent molecules for the simulation study, before the model consistency estimation. Notice that, if the data for the constituent molecules in the models are actually observed, then this process is not necessary. First, the system of differential equations of (3.1)–(3.5) is solved over the time domain. For instance, the solution of the Model A is expressed as follows:

$$\left\{ \begin{aligned} &SLA(t) = SLA(0) \exp(-k_{SN} t), \\ &N(t) = (N(0) - \frac{SLA(0) k_{SN}}{k_{NA} - k_{SN}}) \exp(-k_{NA} t) + \frac{SLA(0) k_{SN}}{k_{NA} - k_{SN}} \exp(-k_{SN} t), \\ &A(t) = \frac{k_{NA} (k_{SN} SLA(0) - k_{NA} N(0) + N(0) k_{SN})}{(k_{NA} - k_{SN})(k_{NA} - k_{AG})} \exp(-k_{NA} t) \\ &\quad + \frac{k_{NA} k_{SN} SLA(0)}{(k_{AG} - k_{SN})(k_{NA} - k_{SN})} \exp(-k_{SN} t) \\ &\quad + (A(0) + \frac{k_{NA} (-k_{SN} SLA(0) + k_{AG} N(0) - N(0) k_{SN})}{(k_{AG} - k_{SN})(k_{NA} - k_{AG})}) \exp(-k_{AG} t), \\ &G(t) = \frac{k_{AG} (-k_{SN} SLA(0) + k_{NA} N(0) - N(0) k_{SN})}{(k_{NA} - k_{AG})(k_{NA} - k_{SN})} \exp(-k_{NA} t) \\ &\quad + \frac{k_{AG} (-k_{NA} + k_{AG}) SLA(0) k_{NA}}{(k_{NA} - k_{AG})(k_{AG} - k_{SN})(k_{NA} - k_{SN})} \exp(-k_{SN} t) \\ &\quad + (-A(0) + \frac{(k_{SN} SLA(0) - k_{AG} N(0) + N(0) k_{SN}) k_{NA}}{(k_{NA} - k_{AG})(k_{AG} - k_{SN})}) \exp(-k_{AG} t) \\ &\quad + SLA(0) + N(0) + A(0) + G(0). \end{aligned} \right. \tag{3.9}$$

In the above equations, we have no information about the actual values of the kinetic parameters and their initial values. Thus, we set them as follows: $k_{SN} = 1$, $k_{NA} = 0.1$, $k_{AG} = 0.5$, $k_{NG} = 0.2$, $k_{SA} = 0.4$, and $k_{SS} = 0.7$ for the kinetic parameters, and $SLA(0) = 10$, $N(0) = 7$, $A(0) = 3$, and $G(0) = 1$ for the initial values. By using the above values, the differential equations of (3.9) are simulated from $t = 0$ to 100 with intervals of 1. Then, we obtain the time series of data for each molecule at 101 sample points. We then numerically estimate the parameters by fitting the equations of (3.7) over the time domain to the above-generated data by the Maple 10 Global Optimization tool (©MapleSoft). In Fig. 4, the sampling data at 101 points and the corresponding equations (fitted curve) are plotted in Model A, together with the given and estimated parameters. Notice that, besides the estimation, all of the parameters in (3.7) can be exactly obtained from the given values for the kinetic parameters and the initial values in (3.9). In the present case, it is natural that the estimated values of the parameters are quite consistent with the given values of the parameters for generating the data.

Consistency Measure. As the first step for the model consistency estimation, we construct a set of polynomials, *comp*, from the algebraic equations of (3.6) for the

models and those of (3.8) for the sampling data. The following equations are *comp* for Model A:

$$\begin{aligned}
 \text{comp} = \{ & k_{SN} - \alpha_{S1A.1}, \\
 & k_{NA} k_{SN} - \alpha_{N.1} \alpha_{N.2}, \\
 & k_{SN} + k_{NA} - \alpha_{N.2} - \alpha_{N.1}, \\
 & N(0) k_{SN} + k_{SN} S L A(0) - \beta_{N.1} \alpha_{N.2} - \beta_{N.2} \alpha_{N.1}, \\
 & k_{AG} + k_{SN} + k_{NA} - \alpha_{A.1} - \alpha_{A.3} - \alpha_{A.2}, \\
 & k_{AG} k_{SN} + k_{AG} k_{NA} + k_{NA} k_{SN} - \alpha_{A.1} \alpha_{A.3} - \alpha_{A.1} \alpha_{A.2} - \alpha_{A.2} \alpha_{A.3}, \\
 & k_{NA} N(0) + A(0) k_{SN} + A(0) k_{NA} - \beta_{A.1} \alpha_{A.3} - \beta_{A.1} \alpha_{A.2} - \beta_{A.2} \alpha_{A.3} - \beta_{A.2} \alpha_{A.1} \\
 & \quad - \beta_{A.3} \alpha_{A.2} - \beta_{A.3} \alpha_{A.1}, \\
 & k_{NA} N(0) k_{SN} + k_{NA} k_{SN} S L A(0) + A(0) k_{NA} k_{SN} - \beta_{A.1} \alpha_{A.2} \alpha_{A.3} - \beta_{A.2} \alpha_{A.1} \alpha_{A.3} \\
 & \quad - \beta_{A.3} \alpha_{A.1} \alpha_{A.2}, \\
 & k_{AG} + k_{SN} + k_{NA} - \alpha_{G.2} - \alpha_{G.1} - \alpha_{G.3}, \\
 & k_{AG} k_{NA} k_{SN} - \alpha_{A.1} \alpha_{A.2} \alpha_{A.3}, k_{AG} k_{NA} k_{SN} - \alpha_{G.1} \alpha_{G.2} \alpha_{G.3}, \\
 & k_{AG} k_{SN} + k_{AG} k_{NA} + k_{NA} k_{SN} - \alpha_{G.1} \alpha_{G.2} - \alpha_{G.2} \alpha_{G.3} - \alpha_{G.1} \alpha_{G.3}, \\
 & k_{AG} k_{NA} N(0) k_{SN} + k_{AG} k_{NA} k_{SN} S L A(0) + G(0) k_{AG} k_{NA} k_{SN} + k_{AG} A(0) k_{NA} k_{SN} \\
 & \quad - \beta_{G.4} \alpha_{G.1} \alpha_{G.2} \alpha_{G.3}, \\
 & G(0) k_{AG} k_{NA} + k_{AG} k_{NA} N(0) + G(0) k_{AG} k_{SN} + k_{AG} A(0) k_{NA} + k_{AG} A(0) k_{SN} \\
 & \quad + G(0) k_{NA} k_{SN} - \beta_{G.3} \alpha_{G.1} \alpha_{G.2} - \beta_{G.1} \alpha_{G.2} \alpha_{G.3} - \beta_{G.2} \alpha_{G.1} \alpha_{G.3} \\
 & \quad - \beta_{G.4} \alpha_{G.1} \alpha_{G.3} - \beta_{G.4} \alpha_{G.2} \alpha_{G.3} - \beta_{G.4} \alpha_{G.1} \alpha_{G.2}, \\
 & G(0) k_{AG} + G(0) k_{NA} + k_{AG} A(0) + G(0) k_{SN} - \beta_{G.2} \alpha_{G.1} - \beta_{G.2} \alpha_{G.3} - \beta_{G.3} \alpha_{G.1} \\
 & \quad - \beta_{G.3} \alpha_{G.2} - \beta_{G.1} \alpha_{G.3} - \beta_{G.4} \alpha_{G.1} - \beta_{G.4} \alpha_{G.2} - \beta_{G.4} \alpha_{G.3} - \beta_{G.1} \alpha_{G.2} \}.
 \end{aligned}$$

In the *comp*, the parameters and the initial values can be expressed as numerical values by the sample data fitting. Thus, only the set of kinetic parameters in the model remains as the unknown parameters in the *comp*. In the following section, we will estimate the kinetic parameters under the constraints in equations (2.4) and (2.6), and will select the model by considering the smallest value of $g(\vec{k})$, the sum-square value of the elements in *comp*.

Model Selection. The model selections by estimating the consistency of the models with the simulated data under the two constraints of equations (2.4) and (2.6) are shown in Table 1. In the first column, the query models, from which the simulated data are generated, are listed, and the models with consistencies that are estimated for the query model are listed in the second column. In the following column, the smallest values of the consistency measure are sorted in ascending order, and the corresponding kinetic measures are listed. As easily seen in this table, the present method has successfully identified the query models. Indeed, all of the models and four of the five models under the two constraints of (2.4) and (2.6) are correctly selected in Table 1, respectively. In addition to the successful selection, the characteristic features for the model selection are observed in the selections by the two constraints. The details of the features are as follows.

As for the selection under the constraint of (2.4), all of the models are clearly selected. By each query model, the corresponding models show the smallest consistency measure (*CM1*) in the constraint of (2.4). For example, when the query model is Model A, the corresponding value for the model consistency for Model A is 1.34×10^{-11} , which

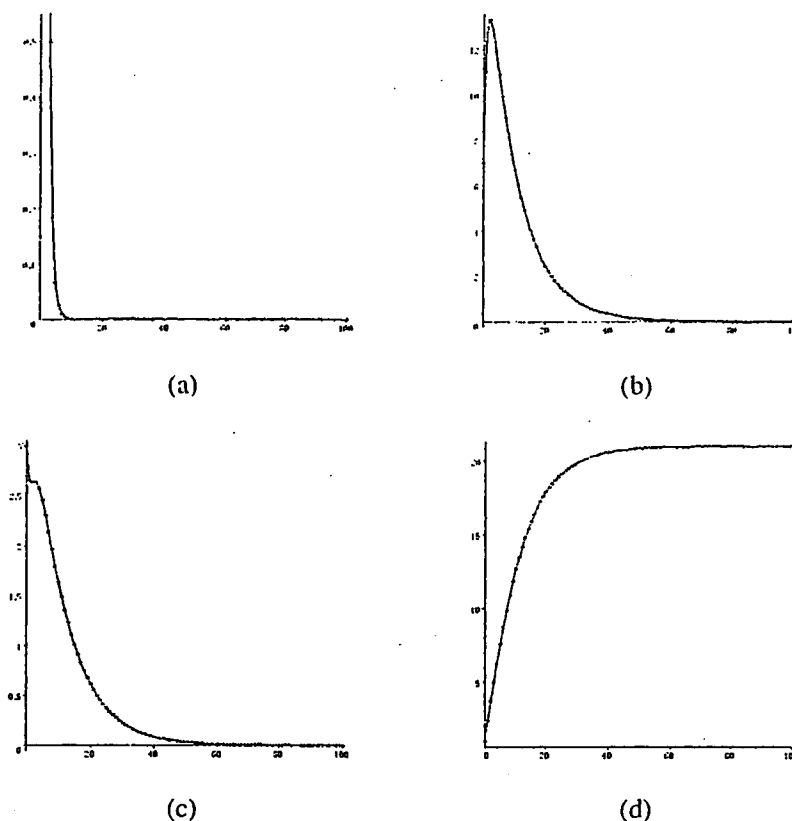


Fig. 4. Sample data for numerical fitting (circles), together with fitted curves (solid lines). The data were generated by numerical calculation from the differential equations (3.9), and the curves were fitted by commercial software (see details in the text). The given and estimated parameters are as follows: $\alpha_{SLA,1}$, 1 (given) and 1.00 (estimated); $\beta_{SLA,1}$, 10 and 10.0; $\alpha_{N,1}$, 1/10 and 0.100; $\alpha_{N,2}$, 1 and 1.00; $\beta_{N,1}$, 163/9 and 18.1; $\beta_{N,2}$, 100/9 and 11.1; $\alpha_{A,1}$, 1/10 and 0.100; $\alpha_{A,2}$, 1/2 and 0.500; $\alpha_{A,3}$, 1 and 1.00; $\beta_{A,1}$, 163/36 and 4.53; $\beta_{A,2}$, $-15/4$ and -3.75 ; $\beta_{A,3}$, 20/9 and 2.22; $\alpha_{G,1}$, 1/10 and 0.100; $\alpha_{G,2}$, 1/2 and 0.500; $\alpha_{G,3}$, 1 and 1.00; $\beta_{G,1}$, $-815/36$ and -22.6 ; $\beta_{G,2}$, 15/4 and 3.75; $\beta_{G,3}$, $-10/9$ and -1.11 ; $\beta_{G,4}$, 21 and 21.0. Each figure corresponds to the four variables (molecules) in the model: (a) *SLA*, (b) *N*, (c) *A*, (d) *G*.

is the smallest among the values of the five models. The magnitude is slightly smaller than 1.36×10^{-11} for Model E. Interestingly, the parameter value for k_{SS} in Model E is estimated to be nearly zero, 1.40×10^{-6} , and when k_{SS} is zero, Model E is identical to Model A. In the remaining models, the parameters cannot be estimated under the constraint of (2.4). In the other query models, the model corresponding to the query model shows the smallest values for the model consistency, and the remaining models show relatively large values or no values, due to the constraint of (2.4). In particular, Model E, in which a cyclic relationship is included, is successfully selected from the other models, especially Model A, which differs from Model E, only in the cyclic part. Furthermore, in all cases, the values of the kinetic parameters are estimated to be equal to the values that are set for the data generation. Thus, the model selection by using the constraint of (2.4) has completely succeeded in all of the models.

Four of the five models are successfully selected under the constraint of (2.6). In the model selection for Model A, Model C is selected. However, Models C, A, E, and B show small values for the consistency measure ($CM2$). Furthermore, three models, Models C, D, and B, become the same form as Model A, by considering the values of the kinetic parameters. Indeed, k_{NG} and $k_{S,A}$ in Model C and $k_{S,A}$ in Model B are estimated to be exactly zero values, and $k_{S,S}$ in Model E is estimated to be a very small value, 1.40×10^{-6} . A similar situation is also found when the query model is Model B. In this case, while the model showing the smallest value is Model B, a similar value is also found in Model C. However, the value of k_{NG} is estimated to be exactly zero, and this indicates that Model C, with the estimated values for kinetic parameters, is the same form as Model B. Thus, the constraint of (2.6) effectively excludes the false relationship between the molecules by estimating the values of the kinetic parameters. As for the model selection for Model E, the small value appears only in the query model, and the relatively large values appear in the other models. In the models with the large values, the $CM2$ values in Models A, B, and C are relatively smaller than the $CM2$ value in Model D. Interestingly, the former models share common chain relationships between S , L , A , N , A , and G with Model E, as seen in Fig 3, while the latter model is a distinctive form from Model E. Even in the inconsistent models, $CM2$ may reflect the similarity of the model form between the query and the estimated models. At any rate, the model selection under the constraint of (2.6) also has succeeded in all of the models.

In summary, the present model selection algorithm shows high performance under the constraints of both (2.4) and (2.6). The constraint of (2.4) focuses on only the selection of a model consistent with the data by a simple algorithm, and the constraint of (2.6) focuses on finer model selection, with the exclusion of false relationships, by a slightly and complicated algorithm. Thus, the algorithm with the constraint of (2.4) is useful to select a model consistent with the data among many candidate models, and that with the constraint of (2.6) is effective to select a model among the candidate models including similar forms.

3.3 Discussion

We have proposed a method for selecting a model that is the most consistent with the data in the present study. In small but distinctive networks, our algorithm has successfully selected the query model, from which the sampling data are generated. The present study partly exploits the previous studies of Cobelli et al. [5, 6] about the relationship between observational parameters and model parameters over the Laplace domain. In these studies, they dealt with the case of differential equations adjoined by a higher dimensional ideal to survey whether the model parameters themselves can be determined uniquely or non-uniquely. In our work, the combination of the transformation of equations over the Laplace domain with the numerical fitting to the observed data enables us to estimate the model's consistency with the data as well as with the values of the kinetic parameters. Although the robustness for data including noise should be further tested, our algorithm is expected to be feasible for actual biological issues regarding the selection of a kinetics model.

The scalability of the present algorithm also remains to be tested. Actually, the present model selection algorithm required several hours for one model. In addition,

Table 1. Model selections. The five models in Fig. 3 were examined for the model selection and the determination of kinetic parameters with the simulated data by the two constraints (see the details in the text). The 'query' and 'estimated' indicate the model from which the simulated data are generated, and the model the consistency of which is estimated by the corresponding query model, respectively.

Query	CM1, under the constraint (2.4)						CM2, under the constraint (2.6)									
	estimated	smallest	k_{SN}	k_{NA}	k_{AG}	k_{NG}	k_{SA}	k_{SS}	estimated	smallest	k_{SN}	k_{NA}	k_{AG}	k_{NG}	k_{SA}	k_{SS}
A	A	1.34×10^{-11}	1.00	0.100	0.500	-	-	-	C	7.66×10^{-12}	1.00	0.100	0.500	0*	0*	-
	E	1.36×10^{-11}	1.00	0.100	0.500	-	-	1.40×10^{-6}	A	1.34×10^{-11}	1.00	0.100	0.500	-	-	-
	D	x	x	-	x	-	x	-	E	1.36×10^{-11}	1.00	0.100	0.500	-	-	1.40×10^{-6}
	B	x	x	x	x	-	x	-	B	1.35×10^{-10}	1.00	0.100	0.500	-	0*	-
	C	x	x	x	x	x	x	-	D	1.68	0*	-	0.435	-	0.0439	-
B	B	4.20×10^{-11}	1.00	0.100	0.500	-	0.400	-	B	4.20×10^{-11}	1.00	0.100	0.500	-	0.400	-
	A	20.1	1.19	0.167	0.637	-	-	-	C	6.44×10^{-11}	1.00	0.100	0.500	0*	0.400	-
	D	x	x	-	x	-	x	-	E	20.1	1.19	0.167	0.637	-	-	0*
	E	x	x	x	x	-	-	x	A	20.1	1.19	0.167	0.637	-	-	-
	C	x	x	x	x	x	x	-	D	1050	0*	-	0.0351	-	1.97	-
C	C	2.78×10^{-9}	1.00	0.100	0.500	0.200	0.400	-	C	2.78×10^{-9}	1.00	0.100	0.500	0.200	0.400	-
	B	0.558	0.994	0.160	0.913	-	0.408	-	B	0.558	0.994	0.160	0.913	-	0.408	-
	A	28.0	1.19	0.213	1.17	-	-	-	D	23.9	0*	-	1.22	-	0.418	-
	D	x	x	-	x	-	x	-	E	28.0	1.19	0.213	1.17	-	-	0*
	E	x	x	x	x	-	-	x	A	28.0	1.19	0.213	1.17	-	-	-
D	D	1.83×10^{-14}	1.00	-	0.500	-	0.400	-	D	1.83×10^{-14}	1.00	-	0.500	-	0.400	-
	A	576	1.02	3.98	0.623	-	-	-	E	358.	1.13	3.63	0.285	-	-	0*
	E	x	x	x	x	-	-	x	B	399.	1.10	3.74	0.395	-	-	0*
	B	x	x	x	x	-	x	-	C	434.	1.18	3.43	0.454	0.528	0*	-
	C	x	x	x	x	x	x	-	A	576.	1.02	3.98	0.623	-	-	-
E	E	9.26×10^{-11}	1.00	0.100	0.500	-	-	0.700	E	9.26×10^{-11}	1.00	0.100	0.500	-	-	0.700
	A	1.46	0.702	0.0564	0.367	-	-	-	C	1.46	0.702	0.0564	0.367	0*	0*	-
	D	x	x	-	x	-	x	-	A	1.46	0.702	0.0564	0.367	-	-	-
	B	x	x	x	x	-	x	-	B	1.46	0.702	0.0564	0.367	-	-	-
	C	x	x	x	x	x	x	-	D	2.57	0*	-	0.258	-	0.0284	-

0*: exact zero value.
 -: no corresponding parameters.
 x: no real positive solutions.

the limit of the nodes and edges in the tested network approximately ranged within 10 edges between 10 nodes. However, the present algorithm over the Laplace domain may overcome the issue of scalability. In a local network within a large-scale network, the relationships of the molecules in the local network with those outside of it are regarded as inputs from the outside, and the variables corresponding to the inputs may easily be eliminated, if the relationships are treated over the Laplace domain. Indeed, we have successfully eliminated the *unnecessary* variables to estimate the parameter values in complex compartmental models for Parkinson's disease by PET measurements [13]. If the *unnecessary* variables in the local network can be eliminated, then the present algorithm can be applied to estimate the model's consistency. Thus, the iteration of the elimination and the consistency estimation may be applicable for the consistency estimation, even in a large-scale network model. Further examinations of the present algorithm for a large-scale network and for noisy data will appear in the near future.

4 Conclusion

In the present model selection, an algebraic manipulation of the differential equations over the Laplace domain, formulated based on the assumption of linear relationships between the variables, is combined with the numerical fitting of the sampling data. The performance of our approach is illustrated with simulated data, in the distinctive forms of models, one of which includes a cyclic relationship hitherto unavailable in previous methods. Although some further examinations of the present method are necessary, especially of the analyzed data and its robustness with noise, the extension of our approach to a large-scale network is promising.

Acknowledgments

H. Y. and K. H. were partly supported by a Grant-in-Aid for Scientific Research on Priority Areas "Systems Genomics" (grant 18016008), by a Grant-in-Aid for Scientific Research (grant 19201039) and by a Grant-in-Aid for Young Scientists (B) (grant 19790881) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT). This study was supported in part by the New Energy and Industrial Technology Development Organization (NEDO) of Japan, by Core Research for Evolutional Science and Technology (CREST), by Program for Improvement of Research Environment for Young Researchers from Special Coordination Funds for Promoting Science and Technology (SCF) commissioned by Japan Science and Technology Agency (JST,MEXT).

References

- [1] Audoly, S., D'Angiò, L., Saccomani, M.P., Cobelli, C.: Global identifiability of linear compartmental models — A computer algebra algorithm. *IEEE Trans. Biomed. Eng.* 45, 36–47 (1998)
- [2] Bisits, A.M., Smith, R., Mesiano, S., Yeo, G., Kwek, K., MacIntyre, D., Chan, E.C.: Inflammatory aetiology of human myometrial activation tested using directed graphs. *PLoS Comput. Biol.* 1, 132–136 (2005)

- [3] Buchberger, B.: An Algorithmic Criterion for the Solvability of a System of Algebraic Equations. In: Buchberger, B., Winkler, F. (eds.) Gröbner Bases and Applications. London Mathematical Society Lecture Notes Series, vol. 251, pp. 535–545. Cambridge University Press, Cambridge (1998)
- [4] Calvano, S.E., Xiao, W., Richards, D.R., Felciano, R.M., Baker, H.V., Cho, R.J., Chen, R.O., Brownstein, B.H., Cobb, J.P., Tschoeke, S.K., Miller-Graziano, C., Moldawer, L.L., Mindrinos, M.N., Davis, R.W., Tompkins, R.G., Lowry, S.F.: Inflammation and Host Response to Injury Large Scale Collab. Res. Program: A network-based analysis of systemic inflammation in humans. *Nature* 437, 1032–1037 (2005)
- [5] Cobelli, C., Foster, D., Toffolo, G.: Tracer Kinetics in Biomedical Research: From Data to Model. Kluwer Academic/Plenum Publishers (2000)
- [6] Cobelli, C., Toffolo, G.: Theoretical aspects and practical strategies for the identification of unidentifiable compartmental systems. ch. 8, pp. 85–91. Pergamon Press, Oxford (1987)
- [7] Hanzon, B., Jibetean, D.: Global minimization of a multivariate polynomial using matrix methods. *Journal of Global Optimization* 27, 1–23 (2003)
- [8] Joreskog, K.G.: A general method for analysis of covariance structures. *Biometrika* 57, 239–251 (1970)
- [9] Meziane, D., Shipley, B.: Direct and Indirect Relationships Between Specific Leaf Area, Leaf Nitrogen and Leaf Gas Exchange. Effects of Irradiance and Nutrient Supply. *Annals of Botany* 88, 915–927 (2001)
- [10] Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, San Francisco (1988)
- [11] Shipley, B.: A new inferential test for path models based on directed acyclic graphs. *Structural Equation Modeling* 7, 206–218 (2000)
- [12] Wright, S.: The method of path coefficients. *Ann. Math. Statist.* 5, 161–215 (1934)
- [13] Yoshida, H., Nakagawa, K., Anai, H., Horimoto, K.: Exact parameter determination for Parkinson’s disease diagnosis with PET using an algebraic approach. In: Anai, H., Horimoto, K., Kutsia, T. (eds.) Algebraic Biology 2007. LNCS, vol. 4545, pp. 110–124. Springer, Heidelberg (2007)