Given two compartments R_1, R_2 that each contain some amount of particles at any time t. They start out with some amount of particles $R_1(0)$ and $R_2(0)$ respectively, s.t. $R_1(0) + R_2(0) = 1$ and they flow out of the compartments via exponential decay with their own rate constants, K_1 and K_2 respectively, s.t. $R_1(t) = R_1(0)e^{-K_1 t}, R_2(t) = R_2(0)e^{-K_2 t}$. They flow into compartment F which initially starts empty ($F(t=0) = 0$), but also flows out via exponential decay as modeled below:

$$\frac{dF}{dt} = -\lambda F(t) + K_1 R_1(t) + K_2 R_2(t)$$

Plugging in for definitions of $R_1(t)$ and $R_2(t)$

$$\frac{dF}{dt} = -\lambda F(t) + K_1 R_1(0)e^{-K_1 t} + K_2 R_2(0)e^{-K_2 t}$$

Bringing the first term over to the left and multiplying by $e^{\lambda t}$

$$\frac{dF}{dt}e^{\lambda t} + \lambda F(t)e^{\lambda t} = K_1 R_1(0)e^{(\lambda - K_1)t} + K_2 R_2(0)e^{(\lambda - K_2)t}$$

Realizing that the left side is equal to the derivative of $F(t)e^{\lambda t}$

$$\frac{d}{dt}(F(t)e^{\lambda t}) = K_1 R_1(0)e^{(\lambda - K_1)t} + K_2 R_2(0)e^{(\lambda - K_2)t}$$

Integrating both sides

$$F(t)e^{\lambda t} = \frac{K_1 R_1(0)e^{(\lambda - K_1)t}}{\lambda - K_1} + \frac{K_2 R_2(0)e^{(\lambda - K_2)t}}{\lambda - K_2} + C$$

Dividing by $e^{\lambda t}$

$$F(t) = \frac{K_1 R_1(0)e^{-K_1 t}}{\lambda - K_1} + \frac{K_2 R_2(0)e^{-K_2 t}}{\lambda - K_2} + Ce^{-\lambda t}$$

Solving for C by saying $F(0) = 0$

$$F(0) = 0 = \frac{K_1 R_1(0)}{\lambda - K_1} + \frac{K_2 R_2(0)}{\lambda - K_2} + C$$

$$C = -\frac{K_1 R_1(0)}{\lambda - K_1} - \frac{K_2 R_2(0)}{\lambda - K_2}$$

Plugging C back in and simplifying

$$F(t) = \frac{K_1 R_1(0)e^{-K_1 t}}{\lambda - K_1} + \frac{K_2 R_2(0)e^{-K_2 t}}{\lambda - K_2} - (\frac{K_1 R_1(0)}{\lambda - K_1} + \frac{K_2 R_2(0)}{\lambda - K_2})e^{-\lambda t}$$

$$F(t) = \frac{K_1 R_1(0)(e^{-K_1 t} - e^{-\lambda t})}{\lambda - K_1} + \frac{K_2 R_2(0)(e^{-K_2 t} - e^{-\lambda t})}{\lambda - K_2}$$

Now, our prediction y at time t is defined as $y = F(t) + R_1(t) + R_2(t)$
We define our prediction of this data point as y By plugging in for the definitions of $R_1(t), R_2(t)$, and $F(t)$

$$y = R_1(0)e^{-K_1 t} + R_2(0)e^{-K_2 t} + \frac{K_1 R_1(0)e^{-K_1 t}}{\lambda - K_1} + \frac{K_2 R_2(0)e^{-K_2 t}}{\lambda - K_2} \\ -(\frac{K_1 R_1(0)}{\lambda - K_1} + \frac{K_2 R_2(0)}{\lambda - K_2})e^{-\lambda t}$$

Combining terms and simplifying

$$y = \frac{\lambda R_1(0)e^{-K_1 t}}{\lambda - K_1} + \frac{\lambda R_2(0)e^{-K_2 t}}{\lambda - K_2} - (\frac{K_1 R_1(0)}{\lambda - K_1} + \frac{K_2 R_2(0)}{\lambda - K_2})e^{-\lambda t}$$

$$y = \frac{R_1(0)(\lambda e^{-K_1 t} - K_1 e^{-\lambda t})}{\lambda - K_1} + \frac{R_2(0)(\lambda e^{-K_2 t} - K_2 e^{-\lambda t})}{\lambda - K_2}$$

To minimize the squared error of prediction y on a single data point d taken at time t we want to minimize

$$(d - y)^2$$

Assuming that λ is fixed so the only variables we are fitting are K_1, K_2, $R_1(0)$, and $R_2(0)$ this is equivalent to minimizing (Note, the d^2 drops out because it is a constant)

$$-2dy + y^2$$

When taking the gradient w.r.t variable z and dropping the 2 this is equivalent to

$$(y - d)\frac{dy}{dz}$$

We then take the gradient of this w.r.t each of the variables.

$$\nabla_{K_1} = (d - y)R_1(0)(\frac{\lambda t e^{-K_1 t} + e^{-\lambda t}}{\lambda - K_1} + \frac{K_1 e^{-\lambda t} - \lambda e^{-K_1 t}}{(\lambda - K_1)^2})$$

$$\nabla_{K_2} = (d - y)R_2(0)(\frac{\lambda t e^{-K_2 t} + e^{-\lambda t}}{\lambda - K_2} + \frac{K_2 e^{-\lambda t} - \lambda e^{-K_2 t}}{(\lambda - K_2)^2})$$

$$\nabla_{R_1(0)} = (y - d)\frac{(\lambda e^{-K_1 t} - K_1 e^{-\lambda t})}{\lambda - K_1}$$

$$\nabla_{R_2(0)} = (y - d)\frac{(\lambda e^{-K_2 t} - K_2 e^{-\lambda t})}{\lambda - K_2}$$

Now, if we add backflow from R_1 and R_2 with rate constant K_3 and the backflow from each sub-compartment is modeled as

$$\frac{dR_1}{dt} = -R_1(t)(K_1 + K_3)$$

$$\frac{dR_2}{dt} = -R_2(t)(K_2 + K_3)$$

Then the equations for how much is in each sub-compartment at any time is given by:

$$R_1 = R_1(0)e^{-t(K_1+K_3)}$$

$$R_2 = R_2(0)e^{-t(K_2+K_3)}$$

Using these updated equations let us re-derive $F(t)$ (Note, many steps are skipped because they are analogous to the steps in the initial derivation of $F(t)$

$$\frac{dF}{dt} = -\lambda F(t) + K_1R_1(t) + K_2R_2(t)$$

Plugging in for definitions of $R_1(t)$ and $R_2(t)$

$$\frac{dF}{dt} = -\lambda F(t) + K_1R_1(0)e^{-t(K_1+k_3)} + K_2R_2(0)e^{-t(K_1+K_3)}$$

After integrating this becomes

$$F(t)e^{\lambda t} = \frac{K_1R_1(0)e^{(\lambda-K_1-K_3)t}}{\lambda - K_1 - K_3} + \frac{K_2R_2(0)e^{(\lambda-K_2-K_3)t}}{\lambda - K_2 - K_3} + C$$

$$C = -\frac{K_1R_1(0)}{\lambda - K_1 - K_3} - \frac{K_2R_2(0)}{\lambda - K_2 - K_3}$$

So

$$F(t) = \frac{K_1R_1(0)e^{-t(K_1+K_3)}}{\lambda - K_1 - K_3} + \frac{K_2R_2(0)e^{-t(K_2+K_3)}}{\lambda - K_2 - K_3}$$
$$-(\frac{K_1R_1(0)}{\lambda - K_1 - K_3} + \frac{K_2R_2(0)}{\lambda - K_2 - K_3})e^{-\lambda t}$$

Again the equation for squared error of prediction y on a single datum d at time t is:

$$(d - y)^2$$

Where $y = F(t) + R_1(t) + R_2(t)$. After combining terms and simplifying

$$y = \frac{R_1(0)(\lambda - K_3)e^{-t(K_1+K_3)} - K_1R_1(0)e^{-\lambda t}}{\lambda - K_1 - K_3}$$
$$+\frac{R_2(0)(\lambda - K_3)e^{-t(K_2+K_3)} - K_2R_2(0)e^{-\lambda t}}{\lambda - K_2 - K_3}$$

Which when minimizing w.r.t. $K_1, K_2, K_3, R_1(0)$, and $R_2(0)$ is equivalent to minimizing

$$-2dy + y^2$$

The gradient with respect to each of the variables is:

$$\nabla_{K_1} = (d-y)R_1(0)(\frac{e^{-\lambda t} + t(\lambda - K_3)e^{-t(K_1+K_3)}}{\lambda - K_1 - K_3} + \frac{K_1 e^{-\lambda t} - (\lambda - K_3)e^{-t(K_1+K_3)}}{(\lambda - K_1 - K_3)^2})$$

$$\nabla_{K_2} = (d-y)R_2(0)(\frac{e^{-\lambda t} + t(\lambda - K_3)e^{-t(K_2+K_3)}}{\lambda - K_2 - K_3} + \frac{K_2 e^{-\lambda t} - (\lambda - K_3)e^{-t(K_2+K_3)}}{(\lambda - K_2 - K_3)^2})$$

$$\nabla_{K_3} = (y-d)(R_1(0)(\frac{(K_3 t - \lambda t - 1)e^{-t(K_1+K_3)}}{\lambda - K_1 - K_3} + \frac{(\lambda - K_3)e^{-t(K_1+K_3)} - K_1 R_1(0)e^{-\lambda t}}{(\lambda - K_1 - K_3)^2}) + R_2(0)(\frac{(K_3 t - \lambda t - 1)e^{-t(K_2+K_3)}}{\lambda - K_2 - K_3} + \frac{(\lambda - K_3)e^{-t(K_2+K_3)} - K_2 R_1(0)e^{-\lambda t}}{(\lambda - K_2 - K_3)^2}))$$

$$\nabla_{R_1(0)} = (y-d)\frac{(\lambda - K_3)e^{-t(K_1+K_3)} - K_1 e^{-\lambda t}}{\lambda - K_1 - K_3}$$

$$\nabla_{R_2(0)} = (y-d)\frac{(\lambda - K_3)e^{-t(K_2+K_3)} - K_1 e^{-\lambda t}}{\lambda - K_2 - K_3}$$