# Reconstruction of Transcription Factor Profiles from Fluorescent Protein Reporter Systems via Dynamic Optimization and Tikhonov Regularization

**Wei Dai and Juergen Hahn**
Dept. of Chemical & Biological Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180

Dept. of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180

**Jia Kang**
Dept. of Chemical Engineering, Texas A&M University, College Station, TX 77843

*DOI 10.1002/aic.14559*
*Published online July 18, 2014 in Wiley Online Library (wileyonlinelibrary.com)*

*This work presents a generally applicable technique for reconstructing transcription factor (TF) profiles from fluorescence microscopy images of green fluorescent protein reporter systems. The approach integrates dynamic optimization and a Tikhonov regularization to avoid over-fitting caused by the highly ill-conditioned structure of this inverse problem. The advantage that the presented approach has over existing methods is that no assumptions are made about the TF profile, the linearity, or lack thereof, of the dynamic model used, and the sampling time of the measurements. Moreover, the method allows to use discretization times for the model different from the measurement sampling times and can also deal with state constraints. The technique has been applied to both simulated and experimental data where the profile of the TFs NF-κB and STAT3 are reconstructed. In both of the case studies, the presented approach exhibits excellent performance while fewer assumptions are needed than for existing techniques.* © 2014 American Institute of Chemical Engineers *AIChE J*, 60: 3754–3761, 2014

*Keywords: green fluorescent protein, dynamic optimization, Tikhonov regularization, dynamic model*

## Introduction

Transcription factors (TFs) play an important role in regulation and control of cellular mechanisms. TFs are activated as part of signal transduction and are key to initiating the transcription event in the nucleus. As such, obtaining information about TF dynamics is important for elucidating the mechanisms involved in signal transduction. There are basically two classes of methods for monitoring changes in the TF concentration. One class directly measures protein concentrations, for example, Western blots or chromatin immunoprecipitation, which can only provide qualitative or semiquantitative data and involves destructive measurement techniques, which prevents the same sample from being used at several time points. Another class of methods is based on inferring the TF concentration from other measurable quantities. This second class of methods includes fluorescent reporter systems, which can provide continuous and noninvasive monitoring of gene expression and transcriptional activity.[1,2] Even though TF concentration cannot be directly monitored using such methods, their profile can be reconstructed from the measured intensity of a fluorescence protein such as the green fluorescent protein (GFP).[3] Models composed of ordinary differential equations (ODEs) have been proposed and used to describe the transcription, translation, and activation of fluorescent proteins.[4,5] Using such a model, it is possible to solve an inverse problem that determines the TF profile from the measured fluorescence dynamics.[6]

As this type of inverse problem is highly ill-conditioned, some form of regularization needs to be applied.[7–9] Two classes of approaches have been investigated to solve the inverse problem from the GFP data[6] using different regularization techniques. One class of techniques are the ones proposed by Finkenstädt et al.,[3] Wang et al.,[10] and Huang et al.,[6] which assumes a certain nature of TF dynamics according to available prior knowledge, that is, the TF profile is described by a predetermined function with unknown parameters, which will be estimated to further characterize the profile. This approach is easy to implement from a numerical point of view, however, it inherently restricts the shape of TF profile, which can potentially be inconsistent with the real profile. An alternate approach proposed by Bansal et al.[7] discretizes the TF profile and the ODE model, and solves the inverse problem via an optimization formulation. This approach is more general as the TF profile will no longer be limited to a specific form,[7] and input constraints can be applied so that the reconstructed profile can lie within a feasible range over time. However, existing work has only involved linear models or certain types of nonlinear models, which can be linearized using a transformation.[7] While this limitation is acceptable for now, it will prevent solution of this inverse problem if more complex mechanisms are added

Correspondence concerning this article should be addressed to J. Hahn at hahnj@rpi.edu.

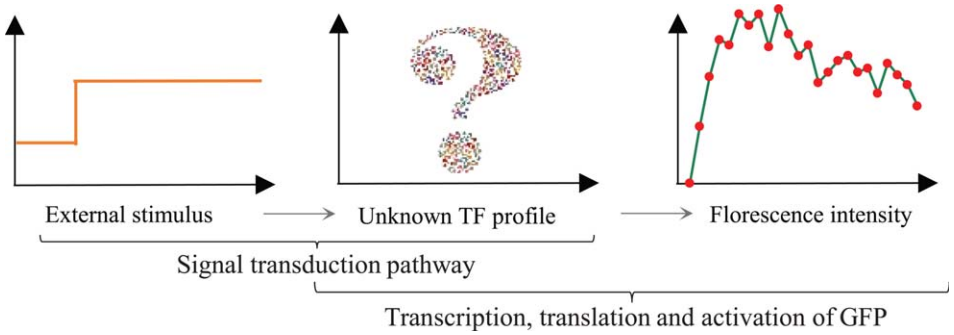


**Figure 1. Scheme for the experimental set up.**

**The external stimulus activates signal transduction which results in dynamics of the transcription factors inside the nucleus. The presence of the activated transcription factor results in proteins being expressed, one of which is a fluorescent protein if a fluorescent reporter system is used. However, the dynamics of interest is the transcription factor concentration and not the measured fluorescence profile. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]**

in the model linking the TF concentration and the measured fluorescence intensity. In addition, the existing technique[7] requires that the fluorescence intensity is measured with a constant sampling time, which may not always be optimal experimentally. Moreover, constraints on state variables cannot be applied, which may result in physical unrealizable solutions such as negative concentrations.

In this work, a more generalized approach is presented that can deal with all the challenges of existing methods. This new technique parameterizes the TF profile as piecewise constant, and discretizes the model via *Radau* collocation on finite elements, thus the inverse problem can be solved as a nonlinear optimization problem. Furthermore, this approach integrates the Tikhonov regularization technique that has been used in previous studies[7] and can overcome over-fitting of the TF profile caused by the highly ill-conditioned structure of the problem. This new method does not place a limitation on the structure of the model, it can directly incorporate state constraints, and it can deal with systems where measurement sampling times are different from the discretization time steps for the model.

The article is structured as follows: preliminaries, including the mathematical model describing transcription, translation, and the activation of GFP and a brief overview of *Radau* collocation on finite elements will be presented first. The technique that integrates dynamic optimization and Tikhonov regularization will be presented in the next section. The following section includes two case studies, which present the reconstruction of NF-$\kappa$B profiles from simulated data and STAT3 profiles from experimental data. The last section contains conclusions.

## Preliminaries

### *Mathematical model*

Stimulation of signal transduction pathway by external stimuli leads ultimately to the activation of TFs in the nucleus. The TF binds to the DNA and initiates the transcription process, which leads to the formation of mRNA. The translation process is initiated by mRNA where the non-fluorescent GFP is formed and then activated to become fluorescent GFP. The presence of activated-GFP results in fluorescence that can be measured. The overall scheme describing this process from the external stimulus to the observed fluorescence intensity is shown in Figure 1. The transcription, translation and the activation of GFP can be described by a mathematical model[4,5]

$$\frac{dm}{dt} = S_m p \frac{C_{\mathrm{TF}}}{c + C_{\mathrm{TF}}} - D_m \tag{1}$$

$$\frac{dn}{dt} = S_n m - D_n n - S_f n \tag{2}$$

$$\frac{df}{dt} = S_f n - D_n f \tag{3}$$

$$I = f/\Delta \tag{4}$$

where the fluorescence intensity, $I$, is the only measured output and the TF concentration in the nucleus, $C_{\mathrm{TF}}$ is the time-dependent input that needs to be reconstructed, $m$ is the mRNA concentration, $n$ is the concentration of non-fluorescent GFP, and $f$ is the concentration of fluorescent GFP. The output of the system $I$ is directly proportional to the concentration of activated GFP, $f$, which is described by Eq. 4. A more detailed description of other individual variables and their (initial) values is given in Tables 1 and 2.

The model (Eqs. 1–4) is a Hammerstein model and it is possible to transform the model into a linear state space model if the Michaelis–Menten kinetics term in Eq. 1, $C_{\mathrm{TF}}/(c + C_{\mathrm{TF}})$, is considered to be a new input. It is worth noting that the presented model is relatively simple and that adding additional mechanisms to the model, such as regulatory mechanisms or using a Hill equation,[11] can result in a truly nonlinear dynamic model.

### *Radau collocation on finite elements*

Discretization of ODEs using collocation points on finite elements is an inherently implicit Runge–Kutta method. The three most commonly used set of collocation points are *Legendre–Gauss* (LG), *Legendre–Gauss–Radau* (LGR, a.k.a. *Radau* collocation), and *Legendre–Gauss–Lobatto* (LGL) points.[12] The precision of LGR, while not quite as good as LG, is very high, and it has added advantage that the end point is one of the abscissas where the function to be integrated, is evaluated.[13] An illustration of the discretization of a continuous time function is shown in Figure 2.

First, a set of finite elements is used to partition the time interval into a number of segments, which usually have the same length. Then multiple collocation points are selected

within each finite element. Next, the continuous function is sampled at each collocation point. By approximating the continuous function with the discrete sampling points, the ODEs (Eq. 5) can be approximated by a set of nonlinear algebraic equations (Eq. 6)[13]

$$\frac{dx}{dt}=f(x,t) \tag{5}$$

$$x_{i,j}=x_{i-1,end}=+h_i\sum_{k\in \mathrm{CP}} f\left(x_{i,k},t_{i,k}\right)\cdot a_{k,j} \tag{6}$$

here $x$ denotes the state vector consisting of entries of the states at different points in time $t$. In the algebraic equations, $i$ is the finite element number, $j$ and $k$ are the collocation point numbers, and CP is the set of collocation points. The word *end* represents the last number of CP, meaning the end point on each finite element. Thus, $x_{i,j}$ denotes the state variable at the $j$th collocation point on the $i$th finite element. The step length of each finite element is $h_i$ (or $h$ for equal-length finite elements), which is specified based on the tradeoff between the accuracy needed for the discretization and the scale of the optimization problem. To obtain a more accurate discretization, $h_i$ can be reduced, or a higher order collocation strategy can be used. *Radau* collocation is widely used for discretizing ODE models since it allows constraints to be set at the end of each element and is a numerically quite stable approach.[14] The coefficient matrix of the *Radau* collocation $\{a_{k,j}\}$ in Eq. 6 is obtained from the roots of a *Legendre* polynomial or linear combinations of a *Legendre* polynomial and its derivatives.[12] In this work, a three-point *Radau* collocation method is used whose coefficient matrix is shown in Eq. 7

$$a=\begin{pmatrix} 0.196815477 & 0.394424315 & 0.376403063 \\ -0.065535426 & 0.292073412 & 0.512485826 \\ 0.023770974 & -0.041548752 & 0.111111111 \end{pmatrix} \tag{7}$$

## Generally Applicable Technique for Computing TF Concentration from Fluorescent Reporter Data

The technique presented in this article is quite general and addresses several of the limitations of existing approaches. The inverse problem is formulated as a dynamic optimization problem and then solved numerically. The resulting optimization problem can be solved with software and hardware that are readily available nowadays on a standard desktop workstation or laptop.

Similar to previous studies,[7] the TF profile is parameterized using piecewise constants, however, the nonlinear ODEs model is directly discretized using *Radau* collocation on finite elements. Although varying-step size finite elements can be used if the system stiffness varies dramatically across the entire time horizon, this work uses a fixed-step three-point *Radau* collocation. In previous studies,[7] the length of the finite element $h$ had to be the same as the step size of the piecewise constant input, and, furthermore, the fluorescence intensity (output) was required to be measured at equally spaced time points. These limitations/assumptions are not required for the method presented here and the TF profile is reconstructed by solving a dynamic optimization problem. To overcome over-fitting of the TF profile, the Tikhonov regularization technique[15] is used in the objective function.

**Table 1. Description and Initial Values of State Variables of the Model Given by Eqs. 1–4**

| State | Description | Initial value (nM) |
|---|---|---|
| $m$ | Concentration of mRNA | 0 |
| $n$ | Concentration of nonfluorescent GFP | 0 |
| $f$ | Concentration of fluorescent GFP | 0 |

**Table 2. The Physical Meanings and Nominal Values of Parameters in the Transcription/Translation/Activation of GFP Model**

| Parameter | Physical meaning | Nominal value |
|---|---|---|
| $S_m$ | Transcription rate | 373 $h^{-1}$ (NF-$\kappa$B)<br>548 $h^{-1}$ (STAT3) |
| $p$ | Amount of available DNA | 5 nM |
| $c$ | Michaelis–Menten constant | 108 nM |
| $D_m$ | Degradation rate of mRNA | 0.45 $h^{-1}$ |
| $S_n$ | Translation rate | 780 $h^{-1}$ |
| $D_n$ | Degradation rate of nonfluorescent GFP | 0.5 $h^{-1}$ |
| $S_f$ | Fluorophore formation rate | 0.347 $h^{-1}$ |
| $\Delta$ | Proportional constant | 25562 nM |
| $C_{\mathrm{TF}}$ | TF concentration in the nucleus | – |
| $I$ | Fluorescence intensity | – |

### *Dynamic optimization including a Tikhonov regularization*

Dynamic optimization refers to a category of optimization problems that address time-varying systems. Generally, these problems seek to maximize or minimize an objective function by determining a group of input profiles that may change over time, while the dynamic model, described by initial-value ODEs, is considered as constraints.[16,17] The reconstruction of TF profiles from GFP data is a dynamic optimization problem. After parameterizing the input TF profile with piecewise constants and discretizing the ODE model with a three-point *Radau* collocation on finite elements, the dynamic optimization problem can be transformed into a general nonlinear programming problem. The problem formulation is shown in Eq. 8.

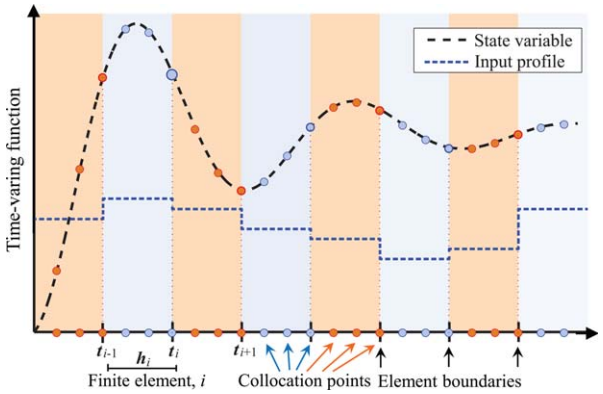


**Figure 2. Illustration of approximating the continuous function with collocation points on finite elements while the input is parameterized with piecewise constants.**

**Here, three collocation points are used within each equal-length finite element. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]**

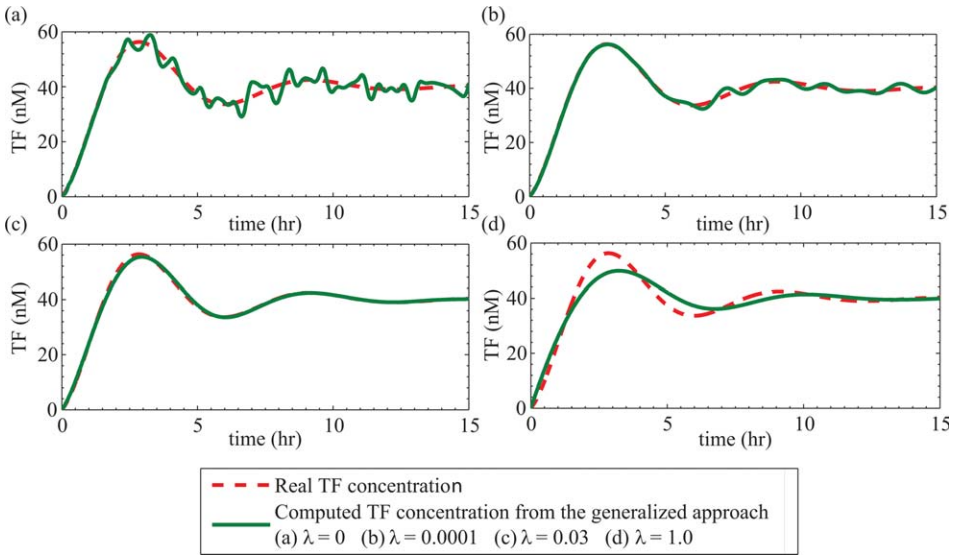


**Figure 3. Reconstruction of TF profile by the presented technique from simulated GFP data without noise.**

**[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]**

**Generally applicable technique based on dynamic optimization including a Tikhonov regularization**

$$\min_{\boldsymbol{C}_{\mathrm{TF}}} \|\boldsymbol{y}\text{-}\hat{\boldsymbol{y}}\|_2^2+\lambda\|\Delta\boldsymbol{C}_{\mathrm{TF}}\|_2^2 \text{ where } \boldsymbol{C}_{\mathrm{TF}} = \{\boldsymbol{C}_{\mathrm{TF}}^r\}, r \in \{0,\ldots,N\}; \Delta\boldsymbol{C}_{\mathrm{TF}}=\{C_{\mathrm{TF}}^r-\boldsymbol{C}_{\mathrm{TF}}^{r-1}\}, r \in \{1,\ldots,N\}$$

$$\text{s.t. } i)\begin{cases} m_{i,j}=x_{i-1,3}+h\sum\limits_{k\in\{1,2,3\}} \dot{m}(m_{i,k},n_{i,k},f_{i,k},\boldsymbol{C}_{\mathrm{TF}}^r)\cdot a_{k,j} \\ n_{i,j}=n_{i-1,3}+h\sum\limits_{k\in\{1,2,3\}} \dot{n}(m_{i,k},n_{i,k},f_{i,k})\cdot a_{k,j} \qquad \text{where, } r=\left\lceil\dfrac{h\cdot i}{h_{\mathrm{TF}}}\right\rceil \quad \text{(Discretized ODEs)} \\ f_{i,j}=f_{i-1,3}+h\sum\limits_{k\in\{1,2,3\}} \dot{f}(m_{i,k},n_{i,k},f_{i,k})\cdot a_{k,j} \end{cases} \tag{8}$$

$$ii) \quad y_{i,j}=f_{i,j}/\Delta, \boldsymbol{y} \text{ is then sampled at measured time points.} \quad \text{(Output sampling)}$$

$$iii) \quad C_{\mathrm{TF}}^r \geq 0, m_{i,j} \geq 0, n_{i,j} \geq 0, f_{i,j} \geq 0, i \in \{1,\ldots N\cdot h_{\mathrm{TF}}/h\} \text{(Non-negativity)}$$

Here the objective function is to minimize the discrepancy between the simulated and experimental data, denoted as $\boldsymbol{y}$ and $\hat{\boldsymbol{y}}$, respectively. A penalty term $\lambda\|\Delta\boldsymbol{C}_{\mathrm{TF}}\|^2$, which is a Tikhonov regularization term, is added to the objective function for regularization purposes. $\boldsymbol{C}_{\mathrm{TF}}$ is a vector consisting of $N+1$ entries where each entry represents the piecewise constant input at step $r$, $r = 0,\ldots, N$. It is common that $C_{\mathrm{TF}}^0$ is fixed at zero since the TF concentration is close to zero when the signal transduction pathway is activated. The step length for parameterizing the input is denoted as $h_{\mathrm{TF}}$. For convenience, $h_{\mathrm{TF}}$ is set to be an integer multiple of $h$, which is the step size for discretizing the model. $\Delta\boldsymbol{C}_{\mathrm{TF}}$ is a vector where each entry is the difference between $\boldsymbol{C}_{\mathrm{TF}}$ at the current step $r$ and that at the previous step $r-1$, where the difference can only be computed for $r \geq 1$.

$\lambda$ is a regularization parameter, which can be tuned to have the problem solved with a stronger (larger $\lambda$) or weaker regularization (smaller $\lambda$). An appropriate $\lambda$ can be selected at the corner of the L-curve,[18–20] which is a plot of the norm of the residual vs. the regularization term for various values of $\lambda$. The two norms vary monotonic with the regularization parameter $\lambda$ with opposite trends and result in an L-shaped curve. The parameter is selected near the corner of this curve to maintain the balance between the norm of the residual and the norm of the solution. This rule of thumb results from the fact that little is gained in terms of minimizing the norm of the solution by increasing the parameter $\lambda$ from the one at the corner value, or with respect to minimizing the residual by decreasing $\lambda$ significantly below the corner value due to the characteristic L-shape of the curve.

The ODE model (Eqs. 1–4) is discretized, resulting in a group of nonlinear algebraic equality constraints that are shown in constraint $i$) of Eq. 8. It is worth noting that since $h_{\mathrm{TF}}$ is an integer multiple of $h$, the input $C_{\mathrm{TF}}^r$ is constant over multiple finite elements, thus being constant in each of the algebraic equations shown in Eq. 8$i$). The index $r$ in each equation can be obtained using the ceiling function denoted by $\lceil\bullet\rceil$ in Eq. 8$i$). In constraint $ii$) of Eq. 8, the output $y_{ij}$ is proportional to the concentration of activated GFP, $f$, and $\Delta$ is the proportional coefficient determined by the specific TF in the experiment. The outputs are calculated at each collocation point of each finite element and then are interpolated at the measured time points. In practice, these

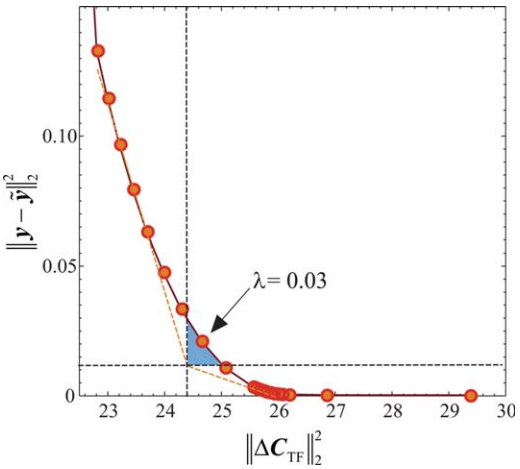


**Figure 4. L-curve for selecting an appropriate value for the regularization parameter for reconstructing the TF profile.**

**[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]**

time points usually lie on the integer multiples of the discretization step. These sampled outputs $\boldsymbol{y}$ are used in the objective function. The non-negativity of the input and states is incorporated in constraint *iii*) of Eq. 8. It is worth noting that additional input and state constraints can be easily applied on the model if it is necessary to do so.

The resulting nonlinear optimization problem is formulated in AMPL[21] and solved by the solver IPOPT,[22] which enjoys excellent convergence properties ($q$-quadratic). A local optimal solution is guaranteed to exist for any specific TF profiles. A global optimal solution could be found by applying global optimization algorithms. While this is not what we have done in the article, the approach is not limited to the optimization algorithm or the shape of TF profiles.

## Case Studies

In this section, the presented procedure for reconstructing TF profiles from GFP data has been applied to both simulated and experimental data. The effects that different approaches and regularization parameters have on the reconstructed TF profiles will also be discussed. The data used in the first case study were created by simulations, thus the real TF profiles are known. The second case study describes the application using experimental GFP data for which the exact nature of the TF profiles are not known, but general trends of the TF profiles are known from other studies.[23–27]

### *Case study 1: simulated data with different levels of Gaussian noise*

The GFP intensity data were generated by simulating the model (Eqs. 1–4) under the assumption that a certain profile for the TF concentration is available. The TF profile is then reconstructed by solving an inverse problem using the dataset via the technique presented in this article and for

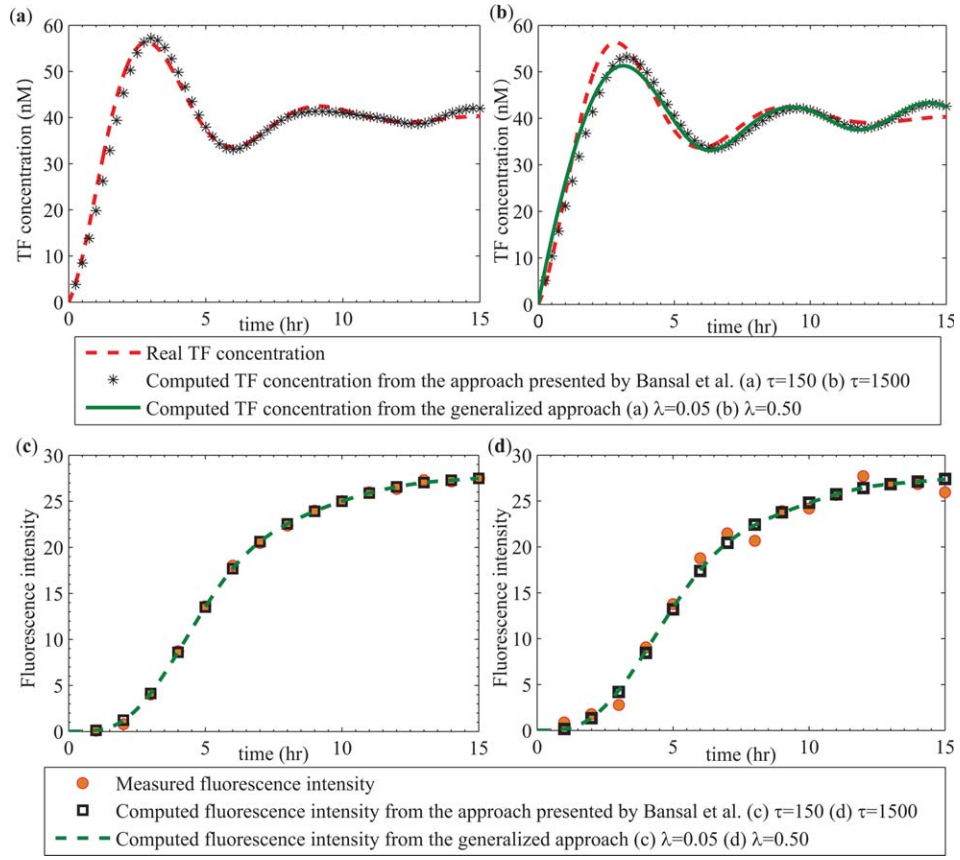


**Figure 5. (a) Reconstruction of TF profile by two different methods with noise $N(0, 0.2)$; (b) Reconstruction of TF profile by two different methods with noise $N(0, 1.0)$; (c) Fitting of the measured fluorescence intensity with noise $N(0, 0.2)$; and (d) Fitting of the measured fluorescence intensity with noise $N(0, 1.0)$.**

**$\tau$ and $\lambda$ are regularization parameters of the two different methods. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]**

**Table 3. Monte Carlo Simulation Results of Datasets Containing Different Levels of Noise**

| | $\tau$ | $\lambda$ | MSE | | Fitting Error | | Bias$^2$ | | Variance | |
|---|---|---|---|---|---|---|---|---|---|---|
| $N$(0,0.2) | 150 | 0.05 | 350 | 331 | 0.535 | 0.545 | 143.7 | 159.4 | 206.6 | 201.6 |
| $N$(0,0.6) | 800 | 0.30 | 987 | 955 | 5.383 | 4.952 | 452.9 | 433.6 | 533.9 | 531.2 |
| $N$(0,0.8) | 1250 | 0.45 | 1303 | 1286 | 9.907 | 9.774 | 605.6 | 557.6 | 696.9 | 697.4 |
| $N$(0,1.0) | 1500 | 0.50 | 1626 | 1512 | 15.45 | 16.51 | 671.5 | 653.1 | 954.6 | 1026 |
| $N$(0,2.0) | 5500 | 2.00 | 3138 | 3359 | 67.32 | 66.30 | 1441 | 1312 | 1696 | 1616 |

The approach presented by Bansal et al. (first) and the technique presented in this work (second) are applied to each dataset.

comparison purposes the approach presented by Bansal et al.[7] While it is not possible to make a comparison for every potential input profile, one commonly found TF profile is given by the following form

$$\boldsymbol{C}_{\mathrm{TF}}(t)=\mathrm{A}(1-e^{-\alpha t}\cos\omega t) \tag{9}$$

where A, $\alpha$, and $\omega$ are known parameters, and $t_0 \le t \le t_f$. This profile represents decaying oscillations which is one of the common dynamics exhibited by TF. For example, the TF dynamics of NF-$\kappa$B for cells continuously stimulated by TNF-$\alpha$ exhibits damped oscillatory behavior, which can be appropriately described by Eq. 9.

At first, the values of A, $\alpha$, and $\omega$ are assumed to be 40, 0.3, and 1.0, respectively. Using this form of TF dynamics, the ODE model was simulated from $t_0 = 0$ to $t_f = 15$ h as most of the TF dynamics will not show significant changes in the fluorescence intensity after this time point. The step size for parameterizing the input, $h_{\mathrm{TF}}$, and for discretizing the model, $h$, must be the same for the technique presented by Bansal et al.[7] The approach presented in this work does not require that $h_{\mathrm{TF}}$ and $h$ have the same value. Nevertheless, the two time steps have been chosen equal to 1 min to appropriately compare the two techniques.

For illustration purpose, the TF profile will first be reconstructed by the presented technique from a simulated dataset where no noise is added (shown in Figure 3). It can be seen that, when the regularization parameter $\lambda$ is set to be zero and no noise is added, the reconstructed TF profile shows significant oscillations. The reason for this is that there are fewer GFP data points than points at which the TF profile is reconstructed which leads to the ill-conditioning of the problem. However, the reconstructed profile becomes smoother as $\lambda$ is increased, and nearly coincides with the real TF profile when $\lambda$ is equal to 0.03. However, when $\lambda$ becomes too large, the reconstructed TF profile becomes too sluggish to follow the dynamics of the real profile as any changes seen in the real profile result in a significant penalty due to the large value of $\lambda$. In other words, the value of $\lambda\|\Delta\boldsymbol{C}_{\mathrm{TF}}\|^2$ makes up a significant portion of the value of the objective function. As the value of the regularization parameter $\lambda$ has a significant impact on the reconstructed TF profile, a systematic procedure to determine appropriate values of $\lambda$ is presented next.

An appropriate value of the regularization parameter $\lambda$ can be selected at the corner of the L-curve, where the $x$ axis is the square of the norm of $\Delta\boldsymbol{C}_{\mathrm{TF}}$, which represents the variation of the TF concentration between each step, and the $y$ axis is the square of the norm of $\boldsymbol{y}-\tilde{\boldsymbol{y}}$, which represents the discrepancy between the simulated data and measured data (Figure 4). Two lines are fitted to the data near each end of the L-curve, and then the candidate regularization parameters can be selected from within the area enclosed by the intersection point, the lines parallel to the axes across the intersection point, and the L-curve.[18–20] There are an infinite number of values of $\lambda$ that lie within this interval and can be chosen, however, all of them will work reasonably well.[28–31]

When no noise is added to the measured data, the presented approach performs well after an appropriate regularization parameter is selected. To create a more realistic dataset, a small amount of additive random Gaussian noise, $N$(0,0.2), is applied to the original dataset. This time, the technique presented by Bansal et al.[7] and the method from this work are used to reconstruct the TF profile. A comparison is shown in Figures 5a, c. It can be seen that both techniques reconstruct the TF profile with good accuracy (shown in Figure 5a) and that the measured intensity data are fitted well (shown in Figure 5c). When a more significant amount of noise $N$(0,1.0) is added to the original measured GFP data, the TF profile can still be reconstructed with a good degree of accuracy using both methods (shown in Figure 5b) and that the measured intensity data are fitted reasonably well (shown in Figure 5d). However, the reconstructed TF profile in Figure 5b is not as good as the one shown in Figure 5a because solving ill-conditioned inverse problems with a significant amount of noise becomes more challenging. It is worth noting that, in Figure 5, $\tau$ and $\lambda$ are regularization parameters for two methods, respectively. No direct comparison can be made between the values of the regularization parameters for these two methods, other than that the regularization parameter values were chosen by the same

**Table 4. Monte Carlo Simulation Results of Datasets Generated by Different Stimulation Profiles**

| | $\alpha = 0.1$ | | $\alpha = 0.2$ | | $\alpha = 0.3$ | | $\alpha = 0.4$ | | $\alpha = 0.5$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $A$=30 | 499 | 504 | 298 | 297 | 232 | 229 | 191 | 188 | 155 | 153 |
| $A$=35 | 618 | 590 | 361 | 374 | 288 | 301 | 234 | 232 | 191 | 186 |
| $A$=40 | 782 | 785 | 435 | 440 | 350 | 331 | 277 | 281 | 230 | 235 |
| $A$=45 | 926 | 964 | 492 | 505 | 395 | 399 | 323 | 337 | 270 | 277 |
| $A$=50 | 1124 | 1171 | 573 | 565 | 447 | 444 | 372 | 359 | 312 | 304 |

Each dataset contains $N$(0, 0.2) noise. The approach presented by Bansal et al. (first) and the technique presented in this work (second) are applied to each dataset. MSEs of the reconstructed TF are compared.

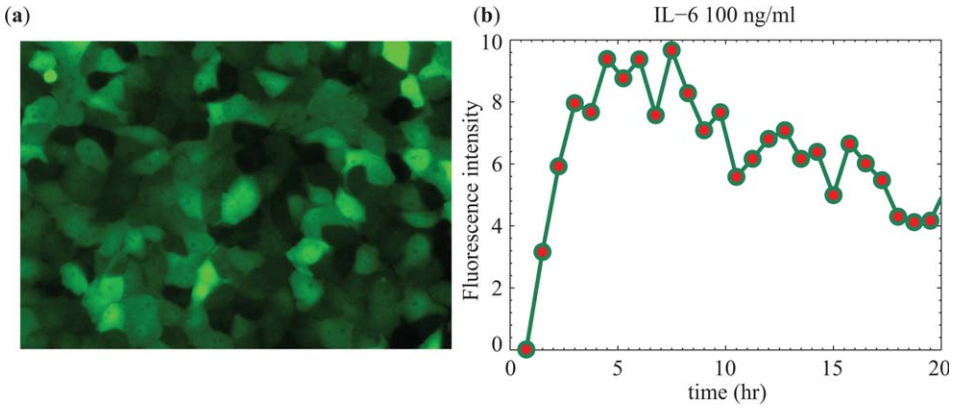


**Figure 6. (a) Sample image obtained from fluorescence microscopy of a GFP reporter system at one point in time and (b) fluorescence intensity profile obtained for IL-6-STAT3 system.**

**[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]**

criteria using an L-curve. In addition, L-curve are different (and have to be different) for each case/method, which explains the different values of the regularization parameters.

The general rule of selecting regularization parameter is that the larger the noise, the larger the regularization parameter has to be chosen, which means that some of the dynamics of the profile will be lost as a result. It can be explained that the larger the measurement noise is, the bigger portion of the norm of residual will take in the objective function, which pushes the corner of L-curve toward the direction of a larger regularization parameter $\lambda$.

To verify that these findings will hold for different parameter values, Monte Carlo simulations are performed next for different levels of noise. The mean squared error, fitting error, squared bias, and variance have been calculated from the reconstructed TF profiles according to Eq. 9, where $\boldsymbol{C}_{\mathrm{TF}}$ is the reconstructed TF concentration, $\bar{\mathbf{C}}_{\mathrm{TF}}$ is its mean value at each time point, and $\hat{\mathbf{C}}_{\mathrm{TF}}$ is the measured TF concentration. For each level of noise, 10,000 datasets with additive random Gaussian noise were used to solve the inverse problem by both the approach presented by Bansal et al. and the technique presented in this work. The results are summarized in Table 3. Monte Carlo simulations are also performed for a variety of stimulation profiles involving different values of the parameter $A$ and $\alpha$ from Eq. 9. Mean square errors (MSEs) of the reconstructed TF profile are used to compare the performance and are summarized in Table 4. All the regularization parameter $\lambda$ used are selected based on L-curve. It can be seen that the results returned by both techniques are comparable, however, the method presented in this work requires fewer limiting assumptions

$$\begin{aligned}\text{Fitting error} &= \|\boldsymbol{y}-\hat{\boldsymbol{y}}\|_2^2\\ \text{MSE} &= \|\mathbf{C}_{\mathrm{TF}}-\hat{\mathbf{C}}_{\mathrm{TF}}\|_2^2\\ \text{Bias}^2 &= \|\hat{\mathbf{C}}_{\mathrm{TF}}-\bar{\mathbf{C}}_{\mathrm{TF}}\|_2^2\\ \text{Variance} &= \|\mathbf{C}_{\mathrm{TF}}-\bar{\mathbf{C}}_{\mathrm{TF}}\|_2^2\end{aligned} \tag{10}$$

### *Case study 2: application to experimental data from Interleukin-6 (IL-6) STAT3 system*

The previous case study illustrated in detail, the effects that different choices of regularization parameters have on the solution of the inverse problem using simulated data under noise conditions. Furthermore, it was shown that the presented technique performs similarly well as an existing method from the literature. In this second case study, the technique presented in this article will be tested on experimental data to determine whether the procedure will return satisfactory results under realistic noise and model uncertainty conditions.

For this case study, experimental data were obtained for the TF STAT3 by continuously stimulating liver cells with 100 ng/mL of IL-6 using a previously developed procedure.[32] The fluorescent microscopy images (Figure 6a) were taken every 45 min for a period of 18 h at multiple positions in the well. The mean fluorescence intensity of the images at each time instant was calculated[4] and is shown in Figure 6b. The shown profile was used to solve the inverse problem to compute the profile for STAT3. It can be seen from these data that the experimental measurements contain a significant amount of noise and a large regularization parameter is needed to suppress over-fitting. Similar to the previous case study, the discretization step size $h$ is set to 1 min. The appropriate regularization parameter $\lambda$ for this data is selected to be 0.25.

It is known that the initial dynamics of the TF STAT3 shows a rapid increase, which results in a pronounced peak of the nuclear STAT3 concentration within the first hour. The reason for this is that cytoplasmic STAT3 is activated and translocates to the nucleus after a few minutes of stimulation with IL-6.[24–26] It can be seen from the results shown in Figure 7 that the STAT3 profile is oscillatory in nature

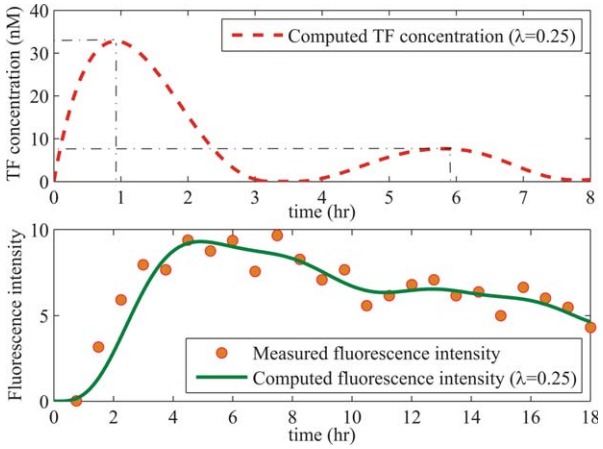


**Figure 7. Reconstruction of TF profile from experimental GFP data obtained for a IL-6-STAT3 reporter system.**

**[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]**

with a large initial peak followed by a smaller peak at approximately 5–6 h. These results are consistent with Western blot data as well as simulation results of the IL-6 signal transduction pathway from the literature.[23,25] Furthermore, the results suggest that the ratio of the peak concentration of STAT3 between the first and the second peak is approximately 5, which is comparable to results reported for simulation studies of the JAK–STAT pathway profile.[27]

## Conclusions

This work presented a technique for computing TF profiles from fluorescence microscopy images of GFP reporter systems. The technique is based on dynamic optimization and a Tikhonov regularization to deal with this ill-conditioned inverse problem. The advantage that the presented technique has over other methods is that no assumptions are made about the TF profile and that it can readily enforce input and state constraints on the model. The method is applied to both simulated and experimental data, and is compared with an existing technique in one of the case studies. The method exhibited excellent performance in both case studies while it requires fewer assumptions than existing techniques.

## Acknowledgments

The authors gratefully acknowledge partial financial support by the National Science Foundation (Grant CBET#0941313) and the American Chemical Society (ACSPRF# 50978-ND9).

## Notation

GFP = green fluorescent protein
IL-6 = Interleukin-6
MSE = mean square error
nM = nano molar
ODE = ordinary differential equation
RE = relative error
STAT3 = signal transducer and activator of transcription 3
TF = transcription factor

## Literature Cited

1. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. Green fluorescent protein as a marker for gene expression. *Science*. 1994; 263(5148):802–805.
2. Van Roessel P, Brand AH. Imaging into the future: visualizing gene expression and protein interactions with fluorescent proteins. *Nat Cell Biol*. 2002;4(1):E15–E20.
3. Finkenstädt B, Heron EA, Komorowski M, Edwards K, Tang S, Harper CV, Davis JRE, White MRH, Millar AJ, Rand DA. Reconstruction of transcriptional dynamics from gene reporter data using differential equations. *Bioinformatics*. 2008;24(24):2901–2907.
4. Huang Z, Senocak F, Jayaraman A, Hahn J. Integrated modeling and experimental approach for determining transcription factor profiles from fluorescent reporter data. *BMC Syst Biol*. 2008;2(1):64.
5. Subramanian S, Srienc F. Quantitative analysis of transient gene expression in mammalian cells using the green fluorescent protein. *J Biotechnol*. 1996;49(1):137–151.
6. Huang Z, Chu Y, Hahn J. Computing transcription factor distribution profiles from green fluorescent protein reporter data. *Chem Eng Sci*. 2012;68(1):340–354.
7. Bansal L, Chu Y, Laird C, Hahn J. Regularization of inverse problems to determine transcription factor profiles from fluorescent reporter systems. *AIChE J*. 2012;58(12):3751–3762.
8. Laird CD, Biegler LT, van Bloemen Waanders BG. Mixed-integer approach for obtaining unique solutions in source inversion of water networks. *J Water Resour Plann Manag*. 2006;132(4):242–251.
9. Laird CD, Biegler LT, van Bloemen Waanders BG, Bartlett RA. Contamination source determination for water networks. *J Water Resour Plann Manag*. 2005;131(2):125–134.
10. Wang X, Errede B, Elston TC. Mathematical analysis and quantification of fluorescent proteins as transcriptional reporters. *Biophys J*. 2008;94(6):2017–2026.
11. Weiss JN. The Hill equation revisited: uses and misuses. *FASEB J*. 1997;11(11):835–841.
12. Garg D, Patterson MA, Francolin C, et al. Direct trajectory optimization and costate estimation of finite-horizon and infinite-horizon optimal control problems using a Radau pseudospectral method. *Comput Optim Appl*. 2011;49(2):335–358.
13. Kameswaran S, Biegler LT. Convergence rates for direct transcription of optimal control problems using collocation at Radau points. *Comput Optim Appl*. 2008;41(1):81–126.
14. Biegler LT, Cervantes AM, Wächter A. Advances in simultaneous strategies for dynamic process optimization. *Chem Eng Sci*. 2002; 57(4):575–593.
15. Golub GH, Hansen PC, O'Leary DP. Tikhonov regularization and total least squares. *SIAM J Matrix Anal Appl*. 1999;21(1): 185–194.
16. Srinivasan B, Palanki S, Bonvin D. Dynamic optimization of batch processes: I. Characterization of the nominal solution. *Comput Chem Eng*. 2003;27(1):1–26.
17. Dai W, Word DP, Hahn J. Modeling and dynamic optimization of fuel-grade ethanol fermentation using fed-batch process. *Control Eng Practice*. 2014;22:231–241.
18. Salvador S, Chan P. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. Paper presented at Tools with Artificial Intelligence. ICTAI 2004. In: 16th IEEE International Conference, Boca Raton, Florida, 2004.
19. Hansen PC. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev*. 1992;34(4):561–580.
20. Johnston PR, Gulrajani RM. Selecting the corner in the L-curve approach to Tikhonov regularization. *IEEE Trans Biomed Eng*. 2000;47(9):1293–1296.
21. Fourer R, Gay D, Kernighan B. AMPL: a modeling language for mathematical programming. Duxbury Press, 2002.
22. Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math Programming*. 2006;106(1):25–57.
23. Fischer P, Lehmann U, Sobota RM, Schmitz J, Niemand C, Linnemann S, Haan S, Behrmann I, Yoshimura A, Johnston JA, Müller-Newen G, Heinrich PC, Schaper F. The role of the inhibitors of interleukin-6 signal transduction SHP2 and SOCS3 for desensitization of interleukin-6 signalling. *Biochem J*. 2004;378:449–460.
24. Kretzschmar A, Dinger M, Henze C, Brocke-Heidrich K, Horn F. Analysis of Stat3 (signal transducer and activator of transcription 3) dimerization by fluorescence resonance energy transfer in living cells. *Biochem J*. 2004;377:289–297.
25. Singh A, Jayaraman A, Hahn J. Modeling regulatory mechanisms in IL-6 signal transduction in hepatocytes. *Biotechnol Bioeng*. 2006; 95(5):850–862.
26. Watanabe K, Saito K, Kinjo M, et al. Molecular dynamics of STAT3 on IL-6 signaling pathway in living cells. *Biochem Biophys Res Commun*. 2004;324(4):1264–1273.
27. Yamada S, Shiono S, Joo A, Yoshimura A. Control mechanism of JAK/STAT signal transduction pathway. *FEBS Lett*. 2003;534(1): 190–196.
28. Calvetti D, Morigi S, Reichel L, Sgallari F. Tikhonov regularization and the L-curve for large discrete ill-posed problems. *J Comput Appl Math*. 2000;123(1):423–446.
29. Hansen PC, O'Leary DP. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J Sci Comput*. 1993;14(6): 1487–1503.
30. Belge M, Kilmer ME, Miller EL. Efficient determination of multiple regularization parameters in a generalized L-curve framework. *Inverse Problems*. 2002;18(4):1161.
31. Hanke M. Limitations of the L-curve method in ill-posed problems. *BIT Numer Math*. 1996;36(2):287–301.
32. Moya C, Huang Z, Cheng P, Jayaraman A, Hahn J. Investigation of IL-6 and IL-10 signalling via mathematical modelling. *IET Syst Biol*. 2011;5(1):15–26.

*Manuscript received Feb. 11, 2014, and revision received July 9, 2014.*