Mathematical optimization approach for estimating the quantum yield distribution of a photochromic reaction in a polymer

Cite as: AIP Advances 7, 015041 (2017); https://doi.org/10.1063/1.4975148
Submitted: 20 October 2016. Accepted: 17 January 2017. Published Online: 31 January 2017

Mirai Tanaka, Takashi Yamashita, Natsuki Sano, Aya Ishigaki, and Tomomichi Suzuki

This paper was selected as an Editor's Pick

ARTICLES YOU MAY BE INTERESTED IN

Density-functional thermochemistry. III. The role of exact exchange
The Journal of Chemical Physics 98, 5648 (1993); https://doi.org/10.1063/1.464913

Spectral curve fitting of dielectric constants
AIP Advances 7, 015042 (2017); https://doi.org/10.1063/1.4975398

Time-resolved optical absorption microspectroscopy of magnetic field sensitive flavin photochemistry
Review of Scientific Instruments 89, 023707 (2018); https://doi.org/10.1063/1.5011693
Mathematical optimization approach for estimating the quantum yield distribution of a photochromic reaction in a polymer

Mirai Tanaka,1,a Takashi Yamashita,2 Natsuki Sano,3 Aya Ishigaki,1 and Tomomichi Suzuki1
1Department of Industrial Administration, Faculty of Science and Technology, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan
2Department of Applied Chemistry, School of Engineering, Tokyo University of Technology, 1404-1 Katakuramachi, Hachioji, Tokyo 192-0982, Japan
3Department of Economics, Management and Information Science, Faculty of Economics, Management and Information Science, Onomichi City University, 1600-2 Hisayamadacho, Onomichi, Hiroshima 722-8506, Japan
(Received 20 October 2016; accepted 17 January 2017; published online 31 January 2017)

The convolution of a series of events is often observed for a variety of phenomena such as the oscillation of a string. A photochemical reaction of a molecule is characterized by a time constant, but materials in the real world contain several molecules with different time constants. Therefore, the kinetics of photochemical reactions of the materials are usually observed with a complexity comparable with those of theoretical kinetic equations. Analysis of the components of the kinetics is quite important for the development of advanced materials. However, with a limited number of exceptions, deconvolution of the observed kinetics has not yet been mathematically solved. In this study, we propose a mathematical optimization approach for estimating the quantum yield distribution of a photochromic reaction in a polymer. In the proposed approach, time-series data of absorbances are acquired and an estimate of the quantum yield distribution is obtained. To estimate the distribution, we solve a mathematical optimization problem to minimize the difference between the input data and a model. This optimization problem involves a differential equation constrained on a functional space as the variable lies in the space of probability distribution functions and the constraints arise from reaction rate equations. This problem can be reformulated as a convex quadratic optimization problem and can be efficiently solved by discretization. Numerical results are also reported here, and they verify the effectiveness of our approach. © 2017 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). [http://dx.doi.org/10.1063/1.4975148]

I. INTRODUCTION

Photochromism is a chemical reaction between two isomers of a molecule. It is induced by light absorption, which reversibly converts one isomer to another.1 One example of a photochromic reaction is the trans to cis isomerization of azobenzene, as shown in Figure 1.

A structural change in the molecule is accompanied by changes in various physical properties such as color, refractive index, polarity, wettability, and electric conductivity. These changes in properties are exploited as functional attributes of the materials and are induced by photo-irradiation as an external switch. Examples of the applications of photochromic molecules are photon-mode optical data storage, color-modulatable glasses, photo-optical materials, photo-mechanical materials, and self-organized structured solar-cell materials.2–7

Corresponding author: mirai@rs.tus.ac.jp.
One of the physical constants of photochromic reactions is quantum yield, which is defined as the ratio of the number of reacted molecules to the number of absorbed photons, and which is usually unique to a molecule. Most organic photo-functional materials comprise photochromic molecules dispersed in a solid-state polymer matrix, where there is a distribution of free volume or the size of the space between polymer chains. Hence, the environment of individual photochromic molecules is different from each other. Therefore, the quantum yield of a photochromic reaction in a solid-state polymer has a large distribution that reflects the inhomogeneity of the polymer matrix, even though the quantum yield of the reaction in a solution is a unique value without a distribution.

There have been several attempts to analyze the multicomponent kinetics. A trivial example is Fourier transformation analysis. One can precisely determine the distribution of the components for a given wave if the observed event is known as the convolution of sine functions. However, most natural events cannot be mathematically solved because their kinetics are mutually coupled with no mutual orthogonal basis set for each.

Representative kinetics for a single molecular reaction is expressed as an exponentially decaying function characterized by a time constant, following the separation of variables that describe the differential equation. To express the inhomogeneity of the time constant, a reaction in viscous media may be analyzed with an extended exponential equation introducing another parameter in addition to the original time constant. Another approach to the analysis of multicomponent kinetics is the deconvolution method, which is usually performed for fluorescence lifetime measurements. One can determine the best-fit solution for a set of parameters by assuming that the fluorescence intensity $I$ can be expressed as

$$I = A_1 \exp \left( -\frac{t}{\tau_1} \right) + A_2 \exp \left( -\frac{t}{\tau_2} \right),$$

where $\tau_1$ and $\tau_2$ are the lifetimes of governing the decaying components of fluorescence; $A_1$ and $A_2$ are the pre-exponential factors, respectively. The problem is that only a limited number of components of the reaction can be expressed on the bases of past methodologies even though there should be a wide distribution of components with an unknown distribution function in practical materials.

We analyzed for the first time the quantum yield distribution of the photochromic reaction of azobenzene in a solid-state polymer to determine several valuable principles of molecular design for high performance materials. However, to simplify the calculations, the analysis was performed assuming that the quantum yield follows a Gaussian distribution. Further progress in the analysis is still limited. Therefore, the development of a novel methodology for solving the distribution function of the components for given data of the convolution of complex phenomena without any assumptions is very important in material science.

In this study, we propose a technique for obtaining the quantum yield distribution of a photochromic reaction based on the time-dependent change in absorbance of a sample. This change in absorbance is described in theory as a convex combination of decaying curves with their respective quantum yields. However, the experimental data also contains noise arising for example from the fluctuation in light intensity, electric noise from detectors, and light scattering at the surface of sample films. Therefore, the experimental data cannot be fitted solely by the convolution of theoretical
equations because of noise, which makes it difficult to estimate the true distribution of the theory-derived equation in the experimental data. Because of this, we propose to solve an optimization problem to obtain a density function that minimizes the difference between the measured and theoretical absorbances corresponding to the density function and to regard it as an estimate. However, it is difficult to solve the optimization problem directly because it is an optimization problem over a functional space and has a differential equation as constraint. In contrast, it is sufficient to obtain a discretized distribution instead of a continuous distribution. Therefore, we discretized the quantum yield distribution to overcome these difficulties. From the discretization, the estimation of the quantum yield density function is reduced to the computation of the quantum yield of each sample point. In addition, we can remove the differential equation constraints by solving the differential equation for each sample point, without loss of equivalence. The resulting optimization model is a convex quadratic optimization problem and is thus easily solved with an existing solver. An optimal solution to the discretized problem is close to that of the original problem if we discretize it sufficiently finely. In general, the solution of the discretized problem converges to the solution of the original problem as mentioned in Still.\textsuperscript{12}

The remainder of this paper is organized as follows: In Section II, we derive a reaction rate equation for photochromism. In Section III, we cover the proposed optimization model. In Section IV, we report on the results of numerical results. Section V is devoted to concluding remarks.

II. REACTION RATE FOR A UNIFORM QUANTUM YIELD

We briefly mention the reaction kinetics of photochromism. First, we introduce the Beer–Lambert law. Then, we derive a reaction rate equation for photochromic molecules. More details about the reaction kinetics of photochromism can be found in the physical chemistry textbook.\textsuperscript{13}

The Beer–Lambert law is a formula of the absorption of light by a material through which light is traveling. Consider light with intensity \( I_0 \) [\( \mu \text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1} \)] irradiating a polymer sample of optical path length \( l \) [cm]. It is known that the intensity \( I \) of the transmitted light decreases exponentially with increasing \( l \). That is, \( I = I_0 e^{-cl} \) or equivalently

\[
\log_{10} \left( \frac{I_0}{I} \right) = bl,
\]

where the constant \( b = a \log_{10}e \) is called the coefficient of absorption. The left-hand side \( \log_{10}(I_0/I) \) is called the absorbance. In addition, we assume there is no interaction between photochromic molecules. Then, the coefficient absorption \( b \) is proportional to the molar concentration \( c \) [mol·L\(^{-1}\)] of photochromic molecules. Let the constant of proportionality be \( \epsilon \) [mol\(^{-1}\)·L·cm\(^{-1}\)]. That is called the molar coefficient of absorption. Then, Equation (1) is reformulated as

\[
\log_{10} \left( \frac{I_0}{I} \right) = \epsilon cl,
\]

which is referred to as the Beer–Lambert law. Let us consider a situation where the molar concentration changes depending with time. Let \( c(t) \) be the molar concentration at time \( t \). Then, the absorbance at time \( t \) is denoted by \( X(t) = \epsilon c(t)l \). From the Beer–Lambert law, we see that the number of photons absorbed by the sample at time \( t \) is \( (I_0 - I)S = I_0S(1 - 10^{-X(t)}) \), where \( S \) is the cross-sectional area of the sample.

Next, we derive a reaction rate equation for photochromic molecules. Here we consider trans-isomers and cis-isomers photochemically changing from one to the other, such as with the photochromism of azobenzene. In this situation, the absorbance \( X(t) \) becomes the sum of that of the trans-isomers and the cis-isomers. Hence, we obtain

\[
X(t) = \epsilon_{\text{trans}}c_{\text{trans}}(t)l + \epsilon_{\text{cis}}c_{\text{cis}}(t)l,
\]

where \( \epsilon_{\text{trans}} \) and \( \epsilon_{\text{cis}} \) denote the molar coefficients of absorption for the trans-isomer and the cis-isomer, respectively; similarly, \( c_{\text{trans}}(t) \) and \( c_{\text{cis}}(t) \) denote the molar concentrations of the trans-isomer and the cis-isomer at time \( t \), respectively. From the results of the Beer–Lambert law, the number of absorbed
photons for the \textit{trans}-isomers is written as
\[ \frac{d}{dt}(10^{-3}c_{\text{trans}}(t)S) = \frac{\epsilon_{\text{trans}}c_{\text{trans}}(t)}{\epsilon_{\text{cis}}c_{\text{cis}}(t) + \epsilon_{\text{trans}}c_{\text{trans}}(t)}I_0S(1 - 10^{-X(t)}). \]

In addition, let \( \phi_{\text{trans} \rightarrow \text{cis}} \) be the quantum yield of the reaction \textit{trans}-isomers to \textit{cis}-isomers and \( \phi_{\text{cis} \rightarrow \text{trans}} \) be that of the reverse reaction. Then, we can obtain the transformation rate from \textit{trans}-isomer to \textit{cis} in [mol \cdot s^{-1}] as follows
\[ \frac{d}{dt}(10^{-3}c_{\text{trans}}(t)S) = \frac{\epsilon_{\text{trans}}c_{\text{trans}}(t)}{\epsilon_{\text{cis}}c_{\text{cis}}(t) + \epsilon_{\text{trans}}c_{\text{trans}}(t)}I_0S(1 - 10^{-X(t)})\phi_{\text{trans} \rightarrow \text{cis}} \]
\[ \quad + \frac{\epsilon_{\text{cis}}c_{\text{cis}}(t)}{\epsilon_{\text{cis}}c_{\text{cis}}(t) + \epsilon_{\text{trans}}c_{\text{trans}}(t)}I_0S(1 - 10^{-X(t)})\phi_{\text{cis} \rightarrow \text{trans}}. \] (3)

The left-hand side is the definition of the transformation rate. The first term on the right-hand side is the decreasing rate of the \textit{trans}-isomers for the forward reaction and the second is the increasing rate of the \textit{trans}-isomers for the backward reaction. In our setting, \( c_{\text{cis}}(0) = 0 \) and hence
\[ c_{\text{trans}}(t) + c_{\text{cis}}(t) = c_{\text{trans}}(0) \] (4)
for any \( t \geq 0 \) hold. Using basic calculus (see Appendix), we obtain the following ordinary differential equation (ODE) of the absorbance \( X(t) \)
\[ \frac{dX(t)}{dt} = 10^3I_0\phi_{\text{trans} \rightarrow \text{cis}}(\epsilon_{\text{trans}} + r_\phi\epsilon_{\text{cis}})(1 - 10^{-X(t)}) \left( \frac{X(0) r_e (1 + r_\phi)}{X(t) (1 + r_\phi r_e)} - 1 \right), \] (5)
where \( r_\phi = \phi_{\text{cis} \rightarrow \text{trans}}/\phi_{\text{trans} \rightarrow \text{cis}} \) and \( r_e = \epsilon_{\text{cis}}/\epsilon_{\text{trans}} \) are substance-specific constants. A solution to Equation (5) depends on the value of the quantum yield \( \phi_{\text{trans} \rightarrow \text{cis}} \) of the forward reaction. In what follows, let \( X(t; \phi) \) be a solution to Equation (5) with \( \phi_{\text{trans} \rightarrow \text{cis}} = \phi \). Hence, if function \( X(t; \phi) \) satisfies Equation (5) with \( \phi_{\text{trans} \rightarrow \text{cis}} = \phi \), we denote the solution as \( \hat{F}(t, X(t; \phi), X(t); \phi) = 0 \).

We plot the numerical solution to Equation (5) using various values of the quantum yield \( \phi \) and the measured absorbance \( y(t) : t \in T \) in Figure 2, where \( T \) is the set of measured times. Details regarding the sample are covered in Section IV. The measured data are plotted as a spiky line. The spikes in the measured data can be regarded as deriving from noise. In contrast, the numerical solutions are drawn as smooth curves. They monotonically decrease and converge to a constant, except at \( \phi = 0 \), the horizontal dotted line. The convergence corresponds to the equilibrium of the forward and backward reactions. Note that the curve of the measured data \( y(t) : t \in T \) does not match any curve from the numerical solution \( X(t; \phi) \). This implies that the quantum yield of the sample is not constant but has a distribution instead.

![FIG. 2. Measured values of the absorbance \( y(t) \) (solid line) and numerical solution \( X(t; \phi) \) to Equation (5) for multiple values of quantum yields (dotted line).](image-url)
III. OPTIMIZATION MODEL FOR ESTIMATING QUANTUM YIELD DISTRIBUTION

We now consider an optimization model to estimate the density function of the quantum yield distribution from the data of the measured absorbance and the reaction rate in Equation (5). Specifically, we propose a method for finding the most suitable density function $\beta$ for the given measured data $(y(t) : t \in \mathcal{T})$ and the parameters in Equation (5). Here, $\beta(\phi)$ denotes the probability density of the existence of molecules whose quantum yield $\phi_{\text{trans} \rightarrow \text{cis}}$ of the forward reaction is $\phi$. In what follows, we assume that $0 \leq \phi_{\text{trans} \rightarrow \text{cis}} \leq 1$. It is equivalent to assuming that no chain reaction occurs during photochromism.

If the quantum yield of the forward reaction takes a significant value $\phi$, we can assume that the measured absorbance $y(t)$ at each time $t \in \mathcal{T}$ is a solution $X(t; \phi)$ of Equation (5) with $\phi_{\text{trans} \rightarrow \text{cis}} = \phi$ with noise $\varepsilon_t$. That is

$$y(t) = X(t; \phi) + \varepsilon_t \quad (t \in \mathcal{T}).$$

Similarly, if the quantum yield has a distribution, we can naturally assume that the measured absorbance $y(t)$ at time $t \in \mathcal{T}$ is an expectation of the solution $X(t; \phi)$ to Equation (5) with the density function $\beta$ and noise $\varepsilon_t$. That is,

$$y(t) = \int_0^1 X(t; \phi) \beta(\phi) d\phi + \varepsilon_t \quad (t \in \mathcal{T}).$$

This model assumes that the measured data $(y(t) : t \in \mathcal{T})$ is a noisy convex combination of $X(t; \phi)$'s with weight $\beta(\phi)$.

Under this assumption, let us consider a way to obtain the most suitable probability distribution function $\beta$ for the measured data. We can model the problem to minimize the sum of squares of the noise intensities $(\varepsilon_t : t \in \mathcal{T})$ under $X(t; \phi)$ to satisfy an initial state condition and Equation (5) as follows:

minimize $\sum_{t \in \mathcal{T}} \left( y(t) - \int_0^1 X(t; \phi) \beta(\phi) d\phi \right)^2$

subject to

$$F(t, X(t; \phi), \dot{X}(t; \phi)) = 0 \quad (t \geq 0, 0 \leq \phi \leq 1),$$

$$X(0; \phi) = X_0 \quad (0 \leq \phi \leq 1),$$

$$\int_0^1 \beta(\phi) d\phi = 1,$$

$$\beta(\phi) \geq 0 \quad (0 \leq \phi \leq 1).$$

This optimization model can be regarded as a constrained least-squares method or a constrained linear regression model by regarding $(y(t) : t \in \mathcal{T})$ as explained variables, $(X(t; \phi) : t \in \mathcal{T})$ as explanatory variables, and $\beta(\phi) (0 \leq \phi \leq 1)$ as coefficients.

The optimization problem expressed by Equation (6) is ill-conditioned. Indeed, for two efficiently close quantum yields, say $\phi$ and $\phi'$, $X(t; \phi)$ is close to $X(t; \phi')$. Hence, there are multiple feasible solutions that attain the same objective value. In the discretized model, to be described below, we see that the matrix of coefficients for the quadratic term in the objective function is ill-conditioned. In terms of linear regression, this is referred to as multicollinearity.

We employ the so-called regularization technique to overcome this problem. In our situation, the $L_1$-regularization is not suitable because we cannot invoke the sparsity in elaborating a solution. Here, we employ the standard $L_2$-regularization. Let $\lambda \geq 0$ be a parameter to define the strength of the regularization. By introducing the $L_2$-regularization term to the objective function of the optimization problem expressed by Equation (6), we obtain the following optimization problem:

minimize $\sum_{t \in \mathcal{T}} \left( y(t) - \int_0^1 X(t; \phi) \beta(\phi) d\phi \right)^2 + \lambda \int_0^1 \beta(\phi)^2 d\phi$

subject to

$$F(t, X(t; \phi), \dot{X}(t; \phi)) = 0 \quad (t \geq 0, 0 \leq \phi \leq 1),$$

$$X(0; \phi) = X_0 \quad (0 \leq \phi \leq 1),$$

$$\int_0^1 \beta(\phi) d\phi = 1,$$

$$\beta(\phi) \geq 0 \quad (0 \leq \phi \leq 1).$$

This optimization model may be regarded as a constrained ridge regression.
The optimization problem expressed by Equation (7) is an ODE-constrained optimization problem on a functional space. Hence, this problem is difficult to solve. Hence, in what follows, by discretizing the problem, we reformulate the model into a convex quadratic programming (QP) model. In more concrete terms, using sampling points \( \phi_0 < \cdots < \phi_N \), we approximate the integral \( \int_0^1 f(\phi) d\phi \) with \( \sum_{i=0}^{N-1} \beta(\phi_i) w_i \), etc. There are multiple means to discretize an integral. Using either Gauss quadrature or Clenshaw-Curtis quadrature,\textsuperscript{14,15} etc. we can achieve an approximation with high-accuracy. In addition, by numerically solving \( F(t, X(t; \phi_i), \dot{X}(t; \phi_i)) = 0 \) with the initial condition \( X(0; \phi_i) = X_0 \) for \( i = 0, \ldots, N \), we can compute a numerical solution \( (X(t; \phi_i) : t \in T) \). By substituting the numerical solution into the optimization problem expressed by Equation (7), we obtain the following optimization model:

\[
\begin{align*}
\text{minimize} & \quad \sum_{t \in T} \left( y(t) - \sum_{i=0}^{N-1} X(t; \phi_i) \beta(\phi_i) w_i \right)^2 + \sum_{i=0}^{N-1} \beta(\phi_i) w_i \\
\text{subject to} & \quad \sum_{i=0}^{N-1} \beta(\phi_i) w_i = 1, \\
& \quad \beta(\phi_i) \geq 0 \quad (i = 0, \ldots, N).
\end{align*}
\]  

As this model is only a convex QP problem, we can easily solve a moderate-sized instance by employing off-the-shelf solvers.\textsuperscript{16-18}

IV. NUMERICAL RESULTS

We now report our numerical results that verify the effectiveness of our approach. Our numerical experiments comprise two parts. One was an experiment to estimate the quantum yield distribution using artificial data generated from a given true quantum yield distribution. We verified that the proposed method estimates with high accuracy the true distribution. The other used empirical data obtained from our measurements. We determined the unknown quantum yield distribution of the sample used in the experiments.

We executed all our experiments in the following environment:

- OS: Mac OS X 10.9.5;
- CPU: Intel Core i7 2.3 GHz;
- RAM: 8 GB;
- A solver for ODEs: ode45 in MATLAB version R2014a;
- A solver for convex QPs: Gurobi Optimizer version 5.6.3.

![FIG. 3. Kullback–Leibler divergence of the true distribution \( \beta \) for the estimate \( \hat{\beta} \) in response to changes in the regularization parameter \( \lambda \).](image)
A. Numerical results for artificial data

We performed the experiments using the following procedure:

1. Choose a true quantum yield distribution $\beta(\phi)$.
2. Choose sampling points $\phi_i$ for $i = 1, \ldots, N$ and numerically solve the reaction rate equation $F(t, X(t; \phi), \dot{X}(t; \phi_i); \phi_i) = 0$ with an appropriate initial condition.
3. Define the set $T$ of (virtual) measuring times, numerically integrate $y(t) = \int_0^1 X(t; \phi) \beta(\phi) d\phi$ for all $t \in T$, and treat them as artificial data of absorbance.
4. For multiple values of the regularization parameter $\lambda$, solve the $L_2$-regularized model, i.e., Equation (8), and compute the difference between the optimal solution $\hat{\beta}$ and the true distribution $\beta$.

We used the following distributions as true quantum yield distributions: Gaussian distributions $N(0.05, 0.01^2)$ and $N(0.05, 0.005^2)$, beta distributions $Be(1, 19)$ and $Be(2, 38)$, and mixed distribution from two Gaussian distributions $\text{Mix}_A = 0.25N(0.02, 0.005^2) + 0.75N(0.06, 0.005^2)$

![FIG. 4. Estimates from artificial data generated from a Gaussian distribution $N(0.05,0.01^2)$.](image1)

![FIG. 5. Estimates from artificial data generated from a Gaussian distribution $N(0.05,0.005^2)$.](image2)
and $\text{Mix}_{\beta} = 0.6N(0.03, 0.005^2) + 0.4N(0.08, 0.005^2)$. We set the set of measured times to $T = \{0, 30, 60, \ldots, 55320\}$ the same as that for the empirical data mentioned in Section IV B. We used the trapezoidal rule in the numerical integration and discretization of integrals. We measured the difference between the optimal solution $\hat{\beta}$ and the true distribution $\beta$ using numerical integration of the Kullback–Leibler divergence\(^{19}\) of $\beta$ for $\hat{\beta}$. That is,

$$D(\beta||\hat{\beta}) = \int_0^1 \beta(\phi) \log \frac{\beta(\phi)}{\hat{\beta}(\phi)} d\phi.$$  

We solved instances of the optimization problem expressed by Equation (8) with multiple values of the regularization parameter $\lambda$. The change in the difference between $\hat{\beta}$ and $\beta$ is shown in Figure 3. We see the following trend for all true distributions: when the regularization parameter $\lambda$ is increased, the difference $D(\beta||\hat{\beta})$ between the optimal solution $\hat{\beta}$ and the true distribution $\beta$ decreases until reaching a minimum near $\lambda = 10^{-9}$; thereafter, it increases.

Next, for each true distribution, we plotted the true distribution (true), an optimal value with $\lambda = 0$ (naive), and an optimal value with an appropriate $\lambda$ that minimized $D(\beta||\hat{\beta})$ (appropriate) in Figures 4–9. In each figure, the optimal solution with $\lambda = 0$ has a number of peaks and is quite different.

![Figure 6](image6.png)

**FIG. 6.** Estimates from artificial data generated from a beta distribution $\text{Be}(1, 19)$.

![Figure 7](image7.png)

**FIG. 7.** Estimates from artificial data generated from a beta distribution $\text{Be}(2, 38)$.
These results imply the following:

- The regularization parameter $\lambda$ has an appropriate value;
- By employing our proposed approach with an appropriate $\lambda$, we can estimate with high accuracy the quantum yield distribution.

### B. Numerical results from empirical data

In the numerical experiments with empirical data, we used the following data measured with the following procedures. Sample films were prepared by bulk polymerization of methyl methacrylate containing azobenzene using AIBN as an initiator. The change in the absorbance of the sample film was recorded by measuring the transmitted light intensity with a photodiode under photoirradiation with 365 nm light from a Xe lamp using a monochromator every 30 s until 55320 s elapsed. The quantum yield of this reaction is known not to exceed unity.
We show the optimal solutions in Figure 10 for when we used the regularization parameters $\lambda = 10^{-9}, 10^{-7}, 10^{-5}$. These values were determined from the numerical results in Section IV A. For each case, we obtained a distribution that has two peaks near $\phi = 0$ and $\phi = 0.055$. The peak at $\phi = 0$ implies the existence of molecules that cannot react because of lack of free volume. In contrast, if a molecule has sufficient free volume, photochromism occurs with quantum yield $\phi = 0.055$.

V. CONCLUDING REMARKS

We proposed an optimization model to estimate the quantum yield distribution of photochromic molecules in a polymer. The problem of finding the most suitable probability distribution function for the quantum yield distribution was modeled as a differential equation constrained optimization problem on a functional space. This can be reduced to a convex quadratic optimization problem by discretizing the space of the quantum yield. Our numerical results imply that

- we can estimate with high accuracy the quantum yield distribution by solving the proposed optimization problem with an appropriate regularization parameter, and
- we can quantitatively analyze empirical data.

In our numerical experiments, we defined an appropriate regularization parameter by executing preparatory experiments. A more objective way is yet to be determined. In terms of material science, our future work includes the development of new devices. To achieve this objective, software based on our approach must be implemented. In addition, a way to solve the optimization problem expressed by Equation (7) without discretization and one to calculate interval estimate of the distribution is an open problem for investigation.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper. They are also grateful to Dr. Yuji Nakatsukasa and Dr. Takayuki Okuno for providing us with useful comments on numerical analysis and semi-infinite programming, respectively.

APPENDIX: DERIVATION OF EQUATION (5) FROM EQUATION (3)

In this appendix, we derive Equation (5) from Equations (2), (3), and (4). By substituting Equation (2) to Equation (3) we obtain
\[
\frac{dc_{\text{trans}}(t)}{dt} = 10^3 I_0 \frac{1 - 10^{-X(t)}}{X(t)} \left( \epsilon_{\text{cis}} c_{\text{cis}}(t) \phi_{\text{cis} \rightarrow \text{trans}} - \epsilon_{\text{trans}} c_{\text{trans}}(t) \phi_{\text{trans} \rightarrow \text{cis}} \right)
\]  

(A1)

In addition, from Equations (2) and (4), we obtain

\[
X(t) = \frac{d}{dt} \left( e_{\text{trans}} c_{\text{trans}}(t) I_0 + e_{\text{cis}} c_{\text{cis}}(t) I_0 - X(t) \right) = 1 - \left( 1 - \frac{e_{\text{cis}}}{e_{\text{trans}}} \right) \frac{c_{\text{cis}}(t)}{c_{\text{trans}}(0)},
\]

where we used \( c_{\text{cis}}(0) = 0 \). Hence, we obtain

\[
X(t) - 1 = \left( \frac{e_{\text{cis}}}{e_{\text{trans}}} - 1 \right) \frac{c_{\text{cis}}(t)}{c_{\text{trans}}(0)} = \left( \frac{e_{\text{cis}}}{e_{\text{cis}}} - 1 \right) \left( 1 - \frac{c_{\text{trans}}(t)}{c_{\text{trans}}(0)} \right).
\]

(A2)

By differentiating it by \( t \), we obtain

\[
\frac{d}{dt} \left( \frac{X(t)}{X(0)} \right) = -\left( \frac{e_{\text{cis}}}{e_{\text{trans}}} - 1 \right) \frac{d c_{\text{trans}}(t)}{d c_{\text{trans}}(0)}.
\]

From this and Equation (A1), we obtain

\[
\frac{dX(t)}{dt} = 10^3 \left( \frac{e_{\text{cis}}}{e_{\text{trans}}} - 1 \right) I_0 (1 - 10^{-X(t)}) \frac{X(0)}{X(t)} \left( \frac{e_{\text{trans}}}{e_{\text{trans}} c_{\text{trans}}(0)} \phi_{\text{trans} \rightarrow \text{cis}} - \epsilon_{\text{cis}} c_{\text{cis}}(t) \phi_{\text{cis} \rightarrow \text{trans}} \right).
\]

Eliminating \( c_{\text{trans}}(t) \) and \( c_{\text{cis}}(t) \) by using Equation (A2), we will see Equation (5).

1. M. Irie, T. Fukaminato, K. Matsuda, and S. Kobatake, Chem. Rev. 114, 12174 (2014).
2. K. Okano, Y. Mikami, Y. Shibata, and T. Yamashita, J. Photopolym. Sci. Technol. 21, 549 (2008).
3. K. Okano, Y. Mikami, and T. Yamashita, Adv. Funct. Mater. 19, 3804 (2009).
4. K. Okano, M. Shinohara, and T. Yamashita, Chem. Eur. J. 15, 3657 (2009).
5. M. Kawamoto, N. Shiga, K. Takaishi, and T. Yamashita, Chem. Commun. 46, 8344 (2010).
6. K. Okano, S. Ogino, M. Kawamoto, and T. Yamashita, Chem. Commun. 47, 11891 (2011).
7. K. Okano, M. Taguchi, M. Fujiki, and T. Yamashita, Angew. Chem. Int. Ed. 50, 12474 (2011).
8. L. Radoszkowicz, J. Pesiado, Y. Erez, E. Nachliel, D. Huppert, and M. Gutman, PCCP 13, 12058 (2011).
9. M. Asano, K. Horie, and T. Yamashita, Polym. Gels Networks 3, 281 (1995).
10. M. Satoh and T. Yamashita, J. Photopolymer. Sci. Technol. 15, 115 (2002).
11. T. Yamashita, K. Ishii, and M. Hasegawa, J. Photopolymer. Sci. Technol. 20, 763 (2007).
12. G. Still, Math. Prog. 91, 53 (2001).
13. P. Atkins and J. C. de Paula, Atkins’s Physical Chemistry, 10th Revised (Oxford University Press, 2014).
14. L. N. Trefethen, SIAM Rev. 50, 67 (2008).
15. L. N. Trefethen, Approximation Theory and Approximation Practice (SIAM, Philadelphia, 2013).
16. FICO, Xpress Optimization Suite, http://www.fico.com/en/Products/DMTools/Pages/FICO-Xpress-Optimization-Suite.aspx.
17. Gurobi Optimization, Inc., Gurobi Optimizer Reference Manual, 2014.
18. IBM, IBM ILOG CPLEX Optimization Studio, http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/.
19. S. Kullback and R. A. Leibler, Ann. Math. Stat. 22, 79 (1951).