IRWIN AND JOAN JACOBS
CENTER FOR COMMUNICATION AND INFORMATION TECHNOLOGIES

Theory of Multiplexing in Fluorescence Microscopy

Marina Alterman, Yoav Y. Schechner and Aryeh Weiss

CCIT Report #729
May 2009

Electronics
Computers
Communications

DEPARTMENT OF ELECTRICAL ENGINEERING
TECHNION - ISRAEL INSTITUTE OF TECHNOLOGY, HAIFA 32000, ISRAEL

CCIT REPORT #729 May 2009

Theory of Multiplexing in Fluorescence Microscopy

MARINA ALTERMAN, YOAV Y. SCHECHNER

Department of Electrical. Eng, Technion - Israel Institute of Technology, Haifa 32000, Israel

and ARYEH WEISS

School of Engineering, Bar-Ilan University, Ramat Gan 52900, Israel

I. INTRODUCTION

Fluorescence microscopy is a powerful tool in biology and biomedical sciences. It is used to study specimens that can fluoresce. In fluorescence, a material is excited by light in one wavelength band and emits light at another band. Both bands are characteristic of the specific material. Microscopic specimens usually yield fluorescence images that are dim and thus suffer from low signal-to-noise ratio (SNR). Apparently, a higher SNR can be achieved either by increasing the intensity of the excitation light source or by using longer exposure times. However, these are often impractical in fluorescence microscopy of living organisms, because the fluorescence intensity is decreased due to photo bleaching. The examination of living organisms is even more difficult since their motion does not permit long exposures. In addition, high intensity illumination can also cause cell damage. We propose a different approach for increasing the SNR in images obtained by a fluorescence microscope. Our approach is based on multiplexing. We are motivated by previous works in the fields of computer vision and photography. They showed that images of dim scenes that were obtained using multiplexed illumination, had higher SNR than single source images. Therefore we propose to multiplex excitation light sources of different wavelengths, as well as different output emission bands in fluorescence microscopy, while imaging multiple fluorescent materials. This way we obtain raw images having higher intensity than a standard image, which is acquired using a single source. Then we decode the desired single source/filter image from the multiplexed images using a computer. We study the limitations of the fluorescence procedure and develop the optimal scheme for multiplexing sources and output filters to achieve the highest quality images. In addition, we propose to use multiplexed images in order to estimate the concentrations (unmixing) of each fluorescent material in the sample more accurately.

II. EXCITATION MULTIPLEXING

A. Case of 3 Excitation Sources and 1 Emission Filter

First we consider a simple case in which the number of excitation wavelengths is 3. An acquired measurement is denoted by a. A measurement i would have been measured, had only one excitation wavelength band been used. Let s denote the index of an excitation source. An estimate of i_s is denoted by $\hat{i}_s$. For example, i_1 is measured if the first excitation wavelength is used.

Now, let two excitation wavelengths bands be used simultaneously, per acquired measurement. For instance, let the second measurement, a_2, be acquired using excitation bands 2 and 3. The fluorescence intensities are additive, therefore,

$$a_2 = i_2 + i_3 \ . \tag{1}$$

Let the third measurement, a_3, be acquired using excitation bands 1 and 3. This measurement is thus

$$a_3 = i_1 + i_3 \ . \tag{2}$$

We construct the vector of individual measurements as,

$$\mathbf{i} = \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} \ . \tag{3}$$

Similarly, the vector of multiplexed measurements is,

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \ . \tag{4}$$

Due to the linear relations between the measurements a and the individual energies i, it is convenient to use a matrix form. There is a multiplexing matrix for excitation, $\mathbf{W}$. For example:

$$\underbrace{\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}}_{\mathbf{a}} = \underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}}_{\mathbf{W}} \underbrace{\begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix}}_{\mathbf{i}} \tag{5}$$

B. General Number of Excitation Sources

Let S denote the number of excitation wavelength sources. The vector of the single source measurements at a single pixel is

$$\mathbf{i} = \begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_{\mathtt{S}} \end{bmatrix} \quad . \tag{6}$$

The elements of $\mathbf{i}$ are denoted by i_s, where s is the index of the excitation source.

We denote by $\mathbf{a}(x, y)$ the vector of multiplexed measurements. The acquired measurements represent fluorescent intensities. Therefore, $\mathbf{i}(x, y)$ and $\mathbf{a}(x, y)$ are related by a linear superposition

$$\mathbf{a}(x, y) = \mathbf{W}\ \mathbf{i}(x, y) \quad , \tag{7}$$

where $\mathbf{W}$ is the multiplexing weighting matrix. Elements of $\mathbf{W}$ are denoted by $w_{m,s}$. Let M be the number of totl measurements. Each element in row $m \in [1, \ldots, M]$ represents the relative power of the excitation source s in measurement m. The power is normalized relatively to its maximum potential value. For example, if $w_{m,s} = 0$, then excitation wavelength s is turned off completely at measurement m; if $w_{m,s} = 1$, then this excitation source is at its maximum power. All excitation sources are turned on with a ratio of their respective maximum power i.e.

$$0 \leq w_{m,s} \leq 1 \quad . \tag{8}$$

III. PHYSICAL MODEL

When a fluorescent specimen is irradiated with a laser source having intensity L_s, the molecules absorb light and become excited. Their return to the ground state is accompanied by emission of light with intensity i. Let s denote the index of an excitation source and d the index of the fluorescent dye. The fluorescent light intensity that is acquired per pixel linearly depends on the following factors:

- $L_s \left[\frac{\mathrm{Watt}}{\mathrm{cm}^2}\right]$ - maximum excitation intensity of source s.
- $t^{\mathtt{exp}}$ $[\mathrm{sec}]$ - exposure time.
- h $[\mathrm{cm}]$ - thickness of the specimen.

- $\alpha_{d,s}$ $[\%]$ - absorbtion coefficinet of fluorophore d under excitation source s. Peak value is normalized to 1.
- e_d $[\%]$ - integral of the emission spectrum (peak value is normalized to 1) of fluorophore d in the emission filter band.
- c_d $\left[\frac{\text{M}}{\text{pixel}}\right]$ - concentration of unbleached fluorophore d.
- Φ_d $[\%]$ - quantum yield for emission of fluorophore d.
- ε_d $[\text{M}^{-1}\text{cm}^{-1}]$ - extinction coefficient of fluorophore d.
- $\tau^{\texttt{optics}}$ $[\%]$ - attenuation constant of the microscope.
- Q^{camera} $[\%]$ - quantum efficiency of the CCD camera.
- $k_{\text{gray}}^{\text{camera}}$ $\left[\frac{\text{graylevel}}{\text{electrons}}\right]$ - digital output of the CCD camera.

Therefore, the fluorescent intensity that is captured per pixel for dye d, while excitation source s is active, can be expressed as

$$i_s^d = k_{\text{gray}}^{\text{camera}}\ Q^{\text{camera}}\ \tau^{\texttt{optics}}\ t^{\text{exp}}\ h\ L_s\ \varepsilon_d\ \alpha_{d,s}\ \Phi_d\ e_d\ c_d \quad . \tag{9}$$

Let D different dyes be present in the specimen. Therefore, using Eq. (9), the fluorescent intensity that is captured per pixel, while excitation source s is active is

$$i_s = k_{\text{gray}}^{\text{camera}}\ Q^{\text{camera}}\ \tau^{\texttt{optics}}\ t^{\text{exp}}\ h\ L_s \sum_{d=1}^{D} \varepsilon_d\ \alpha_{d,s}\ \Phi_d\ e_d\ c_d \quad . \tag{10}$$

The constant values in Eq. (10) can be combined into a single constant and Eq. (10) can be written as

$$i_s = \kappa^{\text{system}}\ L_s \sum_{d=1}^{D} \varepsilon_d\ \alpha_{d,s}\ \Phi_d\ e_d\ c_d \quad , \tag{11}$$

where $\kappa^{\text{system}} = k_{\text{gray}}^{\text{camera}} Q^{\text{camera}} \tau^{\texttt{optics}} t^{\text{exp}} h$.

The physical properties of fluorescent materials that are available for us are sometimes not accurate and therefore it is always preferable to work with normalized properties. Therefore, we define normalized quantities for the source intensities, the quantum efficiency and for the extinction coefficient as follows:

$$L_s^{\text{norm}} \max_{1\ldots S}\{L_s\} = L_s \ ; \quad \Phi_d^{\text{norm}} \max_{1\ldots D}\{\Phi_d\} = \Phi_d \ ; \quad \varepsilon_d^{\text{norm}} \max_{1\ldots D}\{\varepsilon_d\} = \varepsilon_d. \tag{12}$$

By substituting Eq. (12) into Eq. (11), we obtain

$$i_s = \kappa^{\text{system}}\ \kappa^{\max}\ L_s^{\text{norm}} \sum_{d=1}^{D} \varepsilon_d^{\text{norm}}\ \alpha_{d,s}\ \Phi_d^{\text{norm}}\ e_d\ c_d \quad , \tag{13}$$

where $\kappa^{\max} = \max_{1\ldots S}\{L_s\} \max_{1\ldots D}\{\varepsilon_d\} \max_{1\ldots D}\{\Phi_d\}$.

IV. DECODING CONCENTRATIONS - UNMIXING

In this section the purpose is to decode the concentration c_d for each dye d. We can write Eq. (13) in a vector form in the following way:

$$\mathbf{i} = \kappa^{\text{system}}\ \kappa^{\max}\ \mathbf{X}\ \mathbf{c} \quad , \tag{14}$$

where $\mathbf{X}$ is the mixing matrix of size $S \times D$ which has

$$x_{s,d} = L_s^{\text{norm}}\ \varepsilon_d^{\text{norm}}\ \alpha_{d,s}\ \Phi_d^{\text{norm}}\ e_d \tag{15}$$

as its elements. The vector $\mathbf{c}$ of length D is the vector of the concentrations. It has c_d as its elements.

Now, let multiple sources be active in each measurement (multiplexing). Using Eqs. (7) and (14) the vector of multiplexed measurements can be expressed as

$$\mathbf{a} = \kappa^{\text{system}}\ \kappa^{\max}\ \mathbf{W}\ \mathbf{X}\ \mathbf{c} \quad . \tag{16}$$

We wish to decode the concentrations (vector $\mathbf{c}$) from a vector of multiplexed measurements $\mathbf{a}$. Therefore, based on Eq. (16) and the fact that $\mathbf{W}\mathbf{X}$ is generally not square, the estimation of $\mathbf{c}(x, y)$ is computed by

$$\hat{\mathbf{c}}(x, y) = \frac{1}{\kappa^{\text{system}}\ \kappa^{\max}} \left(\mathbf{W}^{\text{x}T}\mathbf{W}^{\text{x}}\right)^{-1} \mathbf{W}^{\text{x}T}\mathbf{a}(x, y) \quad , \tag{17}$$

where

$$\mathbf{W}^{\text{x}} = \mathbf{W}\mathbf{X} \quad . \tag{18}$$

As mentioned before, we do not have very accurate physical properties of the materials we work with. Thus, the estimation of the concentrations up to a constant value is good enough for the unmixing purposes. Therefore, we wish to decode

$$\tilde{\mathbf{c}}(x, y) = \left(\mathbf{W}^{\text{x}T}\mathbf{W}^{\text{x}}\right)^{-1} \mathbf{W}^{\text{x}T}\mathbf{a}(x, y) \quad , \tag{19}$$

where $\tilde{\mathbf{c}} = \hat{\mathbf{c}}\ \kappa^{\text{system}}\ \kappa^{\max}$.

V. OPTIMAL EXCITATION MULTIPLEXING

Let statistically independent additive noise be present in the measurements $\mathbf{a}$. It has zero mean and variance σ^2. Following Eq. (19), this noise propagates to $\tilde{\mathbf{c}}(x, y)$. The mean squared error (MSE) of $\hat{\mathbf{i}}(x, y)$ at each pixel [1], [2] is

$$\mathrm{MSE} = \frac{\sigma^2}{D} \mathrm{trace}\left[(\mathbf{W}^{\mathrm{x}\,T} \mathbf{W}^{\mathrm{x}})^{-1} \right] \quad . \tag{20}$$

In this work, we seek the multiplexing matrix $\mathbf{W}$ that minimizes the MSE.

A. Gradient Descent

In this section, the signal-dependency of the noise is not considered. Therefore, σ^2 is constant. Define

$$\widetilde{\mathrm{MSE}} = \frac{1}{D} \mathrm{trace}\left[(\mathbf{W}^{\mathrm{xT}} \mathbf{W}^{\mathrm{x}})^{-1} \right] \quad . \tag{21}$$

Our optimization problem is equivalent to

$$\widehat{\mathbf{W}} = \arg\min_{\mathbf{W}} \widetilde{\mathrm{MSE}} \quad . \tag{22}$$

Using the gradient descent method [4], $\widetilde{\mathrm{MSE}}$ is iteratively minimized as a function of $\mathbf{W}$. In each iteration n, $\mathbf{W}$ is updated by its gradient $\frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{W}}$:

$$\mathbf{W}_{n+1} = \mathbf{W}_n - \gamma \frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{W}_n} \tag{23}$$

where γ is the parameter controlling the step size. The updated matrix is then projected onto the constraint in Eq. (8).

B. Calculation of the Gradient

To facilitate Eq. (23), we differentiate $\widetilde{\mathrm{MSE}}$ with respect to $\mathbf{W}$. We define $\mathbf{G}$ as the gradient of $\widetilde{\mathrm{MSE}}$ with respect to $\mathbf{W}^{\mathrm{x}}$. Then, we use $\mathbf{G}$ to calculate the partial derivative. From [2],

$$\frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{W}^{\mathrm{x}}} \equiv \mathbf{G} = \frac{-2}{D} (\mathbf{W}^{\mathrm{x}T} \mathbf{W}^{\mathrm{x}} \mathbf{W}^{\mathrm{x}T})^{-1} \quad . \tag{24}$$

We use the following chain rule [5] in order to calculate the partial derivatives:

$$\frac{\partial \widetilde{\mathrm{MSE}}}{\partial w_{m,s}} = \sum_{p=1}^{M} \sum_{d=1}^{D} \frac{\partial \widetilde{\mathrm{MSE}}}{\partial w^{\mathrm{x}}_{p,d}} \cdot \frac{\partial w^{\mathrm{x}}_{p,d}}{\partial w_{m,s}} = \sum_{p=1}^{M} \sum_{d=1}^{D} g_{p,q} \cdot \frac{\partial w^{\mathrm{x}}_{p,d}}{\partial w_{m,s}} \tag{25}$$

where $g_{p,q}$ is an element in $\mathbf{G}$ and $w^{\mathrm{x}}_{p,d}$ is an element in $\mathbf{W}^{\mathrm{x}}$.

Next we explain the computation of the derivatives in Eq. (25). Based on Eq. (18)

$$w^{\mathrm{x}}_{p,d} = \sum_{s=1}^{S} w_{p,s} x_{s,d} \quad . \tag{26}$$

Therefore,

$$\frac{\partial w^{\mathrm{x}}_{p,d}}{\partial w_{m,s}} = \begin{cases} x_{s,d} & p = m \\ 0 & \text{else} \end{cases} \tag{27}$$

Substituting Eq. (27) into Eq. (25)

$$\frac{\partial \widetilde{\mathrm{MSE}}}{\partial w_{m,s}} = \sum_{d=1}^{D} g_{m,d} x_{s,d} \quad . \tag{28}$$

VI. EXCITATION AND EMISSION MULTIPLEXING

A. Case of 3 Excitation Wavelengths and 3 Output Filters

First we consider a special case in which the number of excitation wavelengths is 3, and the number output filters is also 3. An acquired measurement is denoted by a. An energy i would have been measured, had only one excitation wavelength band and one output band-pass filter been used. Let s denote the index of an excitation source and f denote the index of an emission filter band. An estimate of $i_{s,f}$ is denoted by $\hat{i}_{s,f}$. For example, $i_{1,2}$ is the energy measured if the first excitation wavelength is used while the second output band-pass filter is employed.

Now, let two excitation wavelengths bands and two output band-pass filters be used simultaneously, per acquired measurement. For instance, let the fourth measurement, a_4, be acquired using excitation bands 2 and 3, and emission filters 1 and 2. The energies are additive, therefore,

$$a_4 = i_{2,1} + i_{3,1} + i_{2,2} + i_{3,2} \ . \tag{29}$$

Let the fifth measurement, a_5, be acquired using excitation bands 2 and 3, and emission filters 1 and 3. This measurement is thus

$$a_5 = i_{2,1} + i_{3,1} + i_{2,3} + i_{3,3} \ . \tag{30}$$

Our purpose is to estimate the vector of individual energies, $\mathbf{i}$, based on the vector of measurements $\mathbf{a}$. In order to estimate energies for three excitation and three emission bands, nine measurements are needed. We construct the vector of individual energies as,

$$\mathbf{i} = \Bigg[\underbrace{i_{1,1} \quad i_{1,2} \quad i_{1,3}}_{\substack{\text{emission spectrum} \\ \text{under excitation band 1}}} \ \underbrace{i_{2,1} \quad i_{2,2} \quad i_{2,3}}_{\substack{\text{emission spectrum} \\ \text{under excitation band 2}}} \ \underbrace{i_{3,1} \quad i_{3,2} \quad i_{3,3}}_{\substack{\text{emission spectrum} \\ \text{under excitation band 3}}} \Bigg]^T \ , \tag{31}$$

where T denotes transposition. Due to the linear relations between the measurements a and the individual energies i, it is convenient to use a matrix form. There are two multiplexing matrices: one for excitation, $\mathbf{W}^{\texttt{excite}}$, and the other for emission, $\mathbf{W}^{\texttt{emit}}$. First, there is excitation, and therefore $\mathbf{W}^{\texttt{excite}}$ operates on $\mathbf{i}$ before $\mathbf{W}^{\texttt{emit}}$. After the measurements are taken, the vector $\mathbf{i}$ of energies corresponding to the individual excitation and emission bands can be decoded from the vector of measurements $\mathbf{a}$ by inverting the multiplexing matrices.

For each emission band, the energies corresponding to simultaneous active excitations bands are summed. In measurement a_4, (Eq. 29), we first sum the individual energies of the active excitation bands (2 and 3), per each active emission filter. We thus construct the forth and fifth rows of the multiplexing matrix $\mathbf{W}^{\mathtt{excite}}$ as

$$\begin{bmatrix} \vdots \\ i_{2,1} + i_{3,1} \\ i_{2,2} + i_{3,2} \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} & & & & \vdots & & & & \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ & & & & \vdots & & & & \end{bmatrix}}_{\mathbf{W}^{\mathtt{excite}}} \begin{bmatrix} i_{1,1} \\ i_{1,2} \\ i_{1,3} \\ i_{2,1} \\ i_{2,2} \\ i_{2,3} \\ i_{3,1} \\ i_{3,2} \\ i_{3,2} \end{bmatrix} . \tag{32}$$

After excitation, the energies from different active emission filters are summed to obtain a_4. The active emission filters are $f = 1, 2$ and therefore we can construct the forth row of the multiplexing emission matrix, $\mathbf{W}^{\mathtt{emit}}$. This row acts on the result of Eq. (32) to obtain

$$\begin{bmatrix} \vdots \\ a_4 \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ i_{2,1} + i_{3,1} + i_{2,2} + i_{3,2} \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} & & & & \vdots & & & & \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ & & & & \vdots & & & & \end{bmatrix}}_{\mathbf{W}^{\mathtt{emit}}} \begin{bmatrix} \vdots \\ i_{2,1} + i_{3,1} \\ i_{2,2} + i_{3,2} \\ \vdots \end{bmatrix} . \tag{33}$$

To obtain the measurement a_5, (Eq. 30), we only change the emission filters that are activated. Hence, the active emission filters are $f = 1, 3$ while the active excitation sources are unchanged, i.e. $s = 2, 3$. To point to $s = 2, 3$ while $f = 3$, the sixth row of $\mathbf{W}^{\mathtt{excite}}$ is introduced. To mix this results with that of $s = 2, 3$ while $f = 1$, the fifth row of $\mathbf{W}^{\texttt{emit}}$ is introduced:

$$\begin{bmatrix} \vdots \\ a_4 \\ a_5 \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} & & & & \vdots & & & & \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ & & & & \vdots & & & & \end{bmatrix}}_{\mathbf{W}^{\texttt{emit}}} \underbrace{\begin{bmatrix} & & & & \vdots & & & & \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ & & & & \vdots & & & & \end{bmatrix}}_{\mathbf{W}^{\texttt{excite}}} \begin{bmatrix} i_{1,1} \\ i_{1,2} \\ i_{1,3} \\ i_{2,1} \\ i_{2,2} \\ i_{2,3} \\ i_{3,1} \\ i_{3,2} \\ i_{3,3} \end{bmatrix} . \qquad (34)$$

Notice that in Eq. (34), the measurements a_4 and a_5 differ only by the active emission filters, while having the same active excitation sources. In this example, we continue the construction of the multiplexing matrices using two out of three excitation and emission bands in each measurement. Each time, only two of the excitation sources and emission filters are active.

Recall that after the acquisition of $\mathbf{a}$, the vector $\mathbf{i}$ is decoded. This decoding is well-posed when all the permutations of excitations and emissions are covered. The vector $\mathbf{i}$ in Eq. (31) is constructed such that the first three energies are a response to excitation $s = 1$, in three different emission filters (emission spectrum). The next three energies in $\mathbf{i}$ are a response to excitation $s = 2$ in these three emission filters, etc. Therefore in the first three measurements, a_1, a_2 and a_3, we change the emission filters, but use the same excitation. The measurements of the last three components in $\mathbf{a}$ are analogous: but with a different active excitation. The values that are acquired per pixel are

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & & & & & & \\ 1 & 0 & 1 & & 0 & & & 0 & \\ 0 & 1 & 1 & & & & & & \\ & & & 1 & 1 & 0 & & & \\ & 0 & & 1 & 0 & 1 & & 0 & \\ & & & 0 & 1 & 1 & & & \\ & & & & & & 1 & 1 & 0 \\ & 0 & & & 0 & & 1 & 0 & 1 \\ & & & & & & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i_{1,1} \\ i_{1,2} \\ i_{1,3} \\ i_{2,1} \\ i_{2,2} \\ i_{2,3} \\ i_{3,1} \\ i_{3,2} \\ i_{3,3} \end{bmatrix} . \tag{35}$$

The first matrix which operates on $\mathbf{i}$ is $\mathbf{W}^{\texttt{excite}}$. The second matrix is $\mathbf{W}^{\texttt{emit}}$.

B. General Number of Bands

Let S denote the number of excitation wavelength sources and F the number of output emission single band-pass filters. The vector of energy values at a single pixel is

$$\mathbf{i} = \begin{bmatrix} \underbrace{i_{1,1} \quad i_{1,2} \quad \cdots \quad i_{1,F}}_{\text{emission spectrum under excitation band 1}} & \underbrace{i_{2,1} \quad \cdots \quad i_{2,F}}_{\text{emission spectrum under excitation band 2}} & \cdots \quad i_{s,f} \quad \cdots & \underbrace{i_{S,1} \quad \cdots \quad i_{S,F}}_{\text{emission spectrum under excitation band S}} \end{bmatrix}^T . \tag{36}$$

The elements of $\mathbf{i}$ are denoted by $i_{s,f}$, where s is the index of the excitation source and f is the index of the emission filter. In Eq. (36), for excitation source s, the emission spectrum is measured by band-pass filters 1 to F. The consecutive F elements of $\mathbf{i}$ correspond to excitation source $s+1$, and the emission spectrum is represented again by band filters 1 to F, etc.

We denote by $\mathbf{a}(x,y)$ the vector of multiplexed measurements. The number of acquired measurements, which is the length of $\mathbf{a}(x,y)$ is also SF. The acquired measurements represent energies. Therefore, $\mathbf{i}(x,y)$ and $\mathbf{a}(x,y)$ are related by a linear superposition

$$\mathbf{a}(x,y) = \mathbf{W}\ \mathbf{i}(x,y) \quad , \tag{37}$$

where $\mathbf{W}$ is the multiplexing weighting matrix. Based on Eq. (37), the estimation of $\mathbf{i}(x,y)$ is computed by inverting the multiplexing matrix $\mathbf{W}$:

$$\hat{\mathbf{i}}(x,y) = \mathbf{W}^{-1}\ \mathbf{a}(x,y) \quad . \tag{38}$$

Actually, $\mathbf{W}$ is a function of two multiplexing matrices.

$$\mathbf{W} = \mathbf{W}^{\texttt{emit}}\ \mathbf{W}^{\texttt{excite}} \quad , \tag{39}$$

where $\mathbf{W}^{\texttt{emit}}$ is the emission weighting matrix and $\mathbf{W}^{\texttt{excite}}$ is the excitation weighting matrix. Now we present the structure of these weighting matrices.

1) Structure of $\mathbf{W}^{\texttt{excite}}$*:* Let us assume for a moment, that just a single emission band is measured by the system. In this case, the only degree of freedom is the excitation source. Thus the vector of individual energies is

$$\mathbf{i} = \begin{bmatrix} i_{1,1} \\ i_{2,1} \\ \vdots \\ i_{S,1} \end{bmatrix} \quad , \tag{40}$$

of length S. The length of $\mathbf{a}$ is also S. Let $\mathbf{w}^{\text{ex}}$ be an $S \times S$ weighting matrix of different excitation sources. Elements of $\mathbf{w}^{\text{ex}}$ are denoted by $w^{\text{ex}}_{m,s}$. Each element in row $m \in [1, \ldots, S]$ represents the relative power of the excitation source s in measurement m. The power is normalized relatively to its maximum potential value. For example, if $w^{\text{ex}}_{m,s} = 0$, then excitation wavelength s is turned off completely at measurement m; if $w^{\text{ex}}_{m,s} = 1$, then this excitation source is at its maximum power. All excitation sources are turned on with a ratio of their respective maximum power i.e.

$$0 \le w^{\text{ex}}_{m,s} \le 1 \quad \forall m, s \in [1, \ldots, S] \quad . \tag{41}$$

In this case of a single emission band,

$$\mathbf{W} = \mathbf{W}^{\texttt{excite}} = \mathbf{w}^{\text{ex}} \quad . \tag{42}$$

Now, let us consider a more general case, where there are F emission bands (samples of emission spectra). In this general case, the vectors $\mathbf{i}$ and $\mathbf{a}$ are of length SF and $\mathbf{i}$ is arranged as in Eq. (36). Let us examine vector elements corresponding to energies in a specific emission band f but under different excitation sources. As seen in Eq. (36), these elements are interspaced by $F - 1$ unrelated elements. Only the elements having a fixed f and corresponding to different excitations sources need be summed when multiplexing S excitation sources.[1] Let p denote the

---

[1] For now, there is no emission multiplexing.

index of measurement number. For example, let a_p correspond to a measurement using emission filter band $f = 1$:

$$a_p = \sum_{s=1}^{S} w^{\mathtt{ex}}_{m,s} \; i_{s,1} \quad . \tag{43}$$

where the index m is related to p according to:

$$m = \left\lceil \frac{p}{F} \right\rceil \tag{44}$$

Recall that the emission spectrum per excitation is obtained using F measurements, corresponding to F consecutive elements in $\mathbf{i}$. For example, measurement a_{p+1}, which corresponds to emission band $f = 2$, is obtained using the same excitation weights

$$a_{p+1} = \sum_{s=1}^{S} w^{\mathtt{ex}}_{m,s} \; i_{s,2} \tag{45}$$

Using matrix notation for $\mathbf{W}^{\mathtt{excite}}$, this repetition of excitation is expressed by having the elements $w^{\mathtt{ex}}_{m,s}$ interspaced with zeros, and duplicated F times with a shift. This structure is expressed by using the Kronecker product [3]

$$\mathbf{W} = \mathbf{W}^{\mathtt{excite}} = \mathbf{w}^{\mathtt{ex}} \; \otimes \; \mathbf{I}_F \quad , \tag{46}$$

where $\mathbf{I}_F$ is the identity matrix of size $F \times F$ and $\otimes$ denotes the Kronecker product. Eq. (46) yields

$$\mathbf{W}^{\mathtt{excite}} = \begin{bmatrix} w^{\mathtt{ex}}_{1,1}\mathbf{I}_F & \cdots & w^{\mathtt{ex}}_{1,s}\mathbf{I}_F & \cdots & w^{\mathtt{ex}}_{1,S}\mathbf{I}_F \\ \vdots & \ddots & \cdots & \cdots & \vdots \\ w^{\mathtt{ex}}_{m,1}\mathbf{I}_F & \cdots & w^{\mathtt{ex}}_{m,s}\mathbf{I}_F & \cdots & w^{\mathtt{ex}}_{m,S}\mathbf{I}_F \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ w^{\mathtt{ex}}_{S,1}\mathbf{I}_F & \cdots & w^{\mathtt{ex}}_{S,s}\mathbf{I}_F & \cdots & w^{\mathtt{ex}}_{S,S}\mathbf{I}_F \end{bmatrix} \quad . \tag{47}$$

A general element $\{k, l\}$ in $\mathbf{W}^{\mathtt{excite}}$ is denoted by $w^{\mathtt{excite}}_{k,l}$.

2) Structure of $\mathbf{W}^{\mathtt{emit}}$*:* Now we deal with both emission and excitation multiplexing. Let $\mathbf{w}^{\mathtt{em}}_{(s)}$ be an $F \times F$ weighting matrix of different emission filter bands, used under a fixed excitation source s. Elements of $\mathbf{w}^{\mathtt{em}}_{(s)}$ are denoted by $w^{\mathtt{em}}_{(s),r,f}$. Each element in row $r \in [1, \cdots, F]$ represents the transmittance of the corresponding emission filter f in measurement r. For example, if $w^{\mathtt{em}}_{(s),r,f} = 0$, then the filter blocks the energy of band f in measurement r under excitation s; if $w^{\mathtt{em}}_{(s),r,f} = 1$, then all the emission energy of this band is transmitted to the detector. In general,

$$0 \le w^{\mathtt{em}}_{(s),r,f} \le 1 \quad \forall r, f \in [1, \ldots, F] \, , \; \forall s \in [1, \ldots, S] \quad . \tag{48}$$

Only the elements corresponding to a fixed s need to be summed, when multiplexing F emission bands. Consider again the structure of vector $\mathbf{i}$ in Eq. (36). All the elements corresponding to a fixed s reside consecutively in the vector $\mathbf{i}$. Therefore, $\mathbf{W}^{\mathtt{emit}}$ is a block matrix. The diagonal blocks are $\mathbf{w}^{\mathtt{em}}_{(s)}$. The structure of $\mathbf{W}^{\mathtt{emit}}$ is thus:

$$\mathbf{W}^{\mathtt{emit}} = \begin{bmatrix} \mathbf{w}^{\mathtt{em}}_{(1)} & 0 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \cdots & \vdots \\ \vdots & 0 & \mathbf{w}^{\mathtt{em}}_{(s)} & 0 & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 0 & \mathbf{w}^{\mathtt{em}}_{(S)} \end{bmatrix} . \tag{49}$$

A general element $\{h, j\}$ in $\mathbf{W}^{\mathtt{emit}}$ is denoted by $w^{\mathtt{emit}}_{h,j}$.

3) Structure of $\mathbf{W}$*:* Using Eqs. (39,46,47) and (49), we obtain

$$\mathbf{W} = \begin{bmatrix} w^{\mathtt{ex}}_{1,1}\mathbf{w}^{\mathtt{em}}_{(1)} & \cdots & w^{\mathtt{ex}}_{1,s}\mathbf{w}^{\mathtt{em}}_{(1)} & \cdots & w^{\mathtt{ex}}_{1,S}\mathbf{w}^{\mathtt{em}}_{(1)} \\ \vdots & \ddots & \cdots & \cdots & \vdots \\ w^{\mathtt{ex}}_{m,1}\mathbf{w}^{\mathtt{em}}_{(m)} & \cdots & w^{\mathtt{ex}}_{m,s}\mathbf{w}^{\mathtt{em}}_{(m)} & \cdots & w^{\mathtt{ex}}_{m,S}\mathbf{w}^{\mathtt{em}}_{(m)} \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ w^{\mathtt{ex}}_{S,1}\mathbf{w}^{\mathtt{em}}_{(S)} & \cdots & w^{\mathtt{ex}}_{S,s}\mathbf{w}^{\mathtt{em}}_{(S)} & \cdots & w^{\mathtt{ex}}_{S,S}\mathbf{w}^{\mathtt{em}}_{(S)} \end{bmatrix} . \tag{50}$$

A general element $\{p, q\}$ in $\mathbf{W}$ is denoted by $w_{p,q}$ (p is the index of measurement number). Each element in Eq. (50) is

$$w_{p,q} = w^{\mathtt{ex}}_{m,s} w^{\mathtt{em}}_{(m),r,f} \tag{51}$$

such that:

$$m = \left\lceil \frac{p}{F} \right\rceil \quad , \qquad s = \left\lceil \frac{q}{F} \right\rceil \tag{52}$$

$$r = p - (\left\lceil \frac{p}{F} \right\rceil - 1)F \quad , \quad f = q - (\left\lceil \frac{q}{F} \right\rceil - 1)F \tag{53}$$

for all $p, q \in [1, \ldots, SF]$. The inverse formulas are

$$p = r + (m-1)F \quad , \qquad q = f + (s-1)F \tag{54}$$

for all $m, s \in [1, \ldots, S]$ and $r, f \in [1, \ldots, F]$.

VII. OPTIMAL MULTIPLEXING

Similarly to section V let statistically independent additive noise having zero mean and variance σ^2 be present in the measurements $\mathbf{a}$. Following Eq. (38), this noise propagates to $\hat{\mathbf{i}}(x,y)$. The mean squared error (MSE) of $\hat{\mathbf{i}}(x,y)$ at each pixel [1], [2] is

$$\mathrm{MSE} = \frac{\sigma^2}{FS}\mathrm{trace}\left[(\mathbf{W}^T\mathbf{W})^{-1}\right] \quad . \tag{55}$$

Hear, we seek the multiplexing matrix $\mathbf{W}$ that minimizes the MSE.

A. Gradient Descent

In this section, the signal-dependency of the noise is not considered. Therefore, σ^2 is constant. Define

$$\widetilde{\mathrm{MSE}} = \frac{1}{FS}\mathrm{trace}\left[(\mathbf{W}^T\mathbf{W})^{-1}\right] \quad . \tag{56}$$

Our optimization problem is equivalent to

$$\widehat{\mathbf{W}} = \arg\min_{\mathbf{W}} \widetilde{\mathrm{MSE}} \quad . \tag{57}$$

Using the gradient descent method [4], $\widetilde{\mathrm{MSE}}$ is iteratively minimized as a function of $\mathbf{W}^{\mathtt{excite}}$ and $\mathbf{W}^{\mathtt{emit}}$. In each iteration n, $\mathbf{W}^{\mathtt{excite}}$ and $\mathbf{W}^{\mathtt{emit}}$ are updated by their gradients $\frac{\partial\widetilde{\mathrm{MSE}}}{\partial\mathbf{W}^{\mathtt{excite}}}$ and $\frac{\partial\widetilde{\mathrm{MSE}}}{\partial\mathbf{W}^{\mathtt{emit}}}$:

$$\mathbf{W}^{\mathtt{excite}}_{n+1} = \mathbf{W}^{\mathtt{excite}}_{n} - \gamma\frac{\partial\widetilde{\mathrm{MSE}}}{\partial\mathbf{W}^{\mathtt{excite}}_{n}} \tag{58}$$

$$\mathbf{W}^{\mathtt{emit}}_{n+1} = \mathbf{W}^{\mathtt{emit}}_{n} - \gamma\frac{\partial\widetilde{\mathrm{MSE}}}{\partial\mathbf{W}^{\mathtt{emit}}_{n}} \quad , \tag{59}$$

where γ is the parameter controlling the step size. The updated matrices are then projected onto the constraints in Eqs. (41) and (48).

B. Calculation of Gradients

To facilitate Eqs. (58) and (59), we differentiate $\widetilde{\mathrm{MSE}}$ with respect to $\mathbf{W}^{\mathtt{excite}}$ and with respect to $\mathbf{W}^{\mathtt{emit}}$, separately. We define $\mathbf{G}$ as the gradient of $\widetilde{\mathrm{MSE}}$ with respect to $\mathbf{W}$. Then, we use $\mathbf{G}$ to calculate the partial derivatives. From [2],

$$\frac{\partial\widetilde{\mathrm{MSE}}}{\partial\mathbf{W}} \equiv \mathbf{G} = \frac{-2}{FS}(\mathbf{W}^T\mathbf{W}\mathbf{W}^T)^{-1} \quad . \tag{60}$$

We use the following chain rule [5] in order to calculate the partial derivatives:

$$\frac{\partial \widetilde{\mathrm{MSE}}}{\partial w^{\texttt{excite}}_{k,l}} = \sum_p \sum_q \frac{\partial \widetilde{\mathrm{MSE}}}{\partial w_{p,q}} \cdot \frac{\partial w_{p,q}}{\partial w^{\texttt{excite}}_{k,l}} = \sum_p \sum_q g_{p,q} \cdot \frac{\partial w_{p,q}}{\partial w^{\texttt{excite}}_{k,l}} \tag{61}$$

$$\frac{\partial \widetilde{\mathrm{MSE}}}{\partial w^{\texttt{emit}}_{h,j}} = \sum_p \sum_q \frac{\partial \widetilde{\mathrm{MSE}}}{\partial w_{p,q}} \cdot \frac{\partial w_{p,q}}{\partial w^{\texttt{emit}}_{h,j}} = \sum_p \sum_q g_{p,q} \cdot \frac{\partial w_{p,q}}{\partial w^{\texttt{emit}}_{h,j}} \quad , \tag{62}$$

where $g_{p,q}$ is an element in $\mathbf{G}$. In sections VII-B.1 and VII-B.2 we explain in detail the computation of the derivatives in Eqs. (61) and (62).

1) Calculation of $\frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{W}^{\texttt{excite}}}$: Let us examine the structure of $\mathbf{W}^{\texttt{excite}}$ in Eq. (47). Notice that the actual number of degree of freedom is less than $SF \times SF$. The reason is that there are $S \times S$ elements that are repeated F times each, while the other elements of $\mathbf{W}^{\texttt{excite}}$ have null value. Therefore, the formulation in Eq. (61) is both inefficient and redundant. We exploit the sparsity and redundancy in the structure of $\mathbf{W}^{\texttt{excite}}$: we can differentiate $w_{p,q}$ in Eq. (61) only with respect to the elements $w^{\texttt{ex}}_{m,s}$. We divide $\mathbf{G}$ into blocks of $F \times F$ sub-matrices $\widetilde{\mathbf{G}}^{\texttt{ex}}_{m,s}$:

$$\mathbf{G} = \begin{bmatrix} \widetilde{\mathbf{G}}^{\texttt{ex}}_{1,1} & \cdots & \widetilde{\mathbf{G}}^{\texttt{ex}}_{1,S} \\ \vdots & \ddots & \vdots \\ \widetilde{\mathbf{G}}^{\texttt{ex}}_{S,1} & \cdots & \widetilde{\mathbf{G}}^{\texttt{ex}}_{S,S} \end{bmatrix} \quad . \tag{63}$$

Eq. (61) is a summation over element-wise multiplications. Based on Eq. (50), this summation (Eq. 61) can be expressed as,

$$\frac{\partial \widetilde{\mathrm{MSE}}}{\partial w^{\texttt{ex}}_{m,s}} = \mathrm{trace}\left[\widetilde{\mathbf{G}}^{\texttt{ex}}_{m,s} (\mathbf{w}^{\texttt{em}}_{(s)})^T\right] \quad , \tag{64}$$

for all $m, s \in [1, \ldots S,]$. In view of Eqs. (47) and (61), the gradient of $\widetilde{\mathrm{MSE}}$ with respect to $\mathbf{W}^{\texttt{excite}}$ can be represented as:

$$\frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{W}^{\texttt{excite}}} = \begin{bmatrix} \frac{\partial \widetilde{\mathrm{MSE}}}{\partial w^{\texttt{ex}}_{1,1}} \mathbf{I}_F & \cdots & \frac{\partial \widetilde{\mathrm{MSE}}}{\partial w^{\texttt{ex}}_{1,S}} \mathbf{I}_F \\ \vdots & \ddots & \vdots \\ \frac{\partial \widetilde{\mathrm{MSE}}}{\partial w^{\texttt{ex}}_{S,1}} \mathbf{I}_F & \cdots & \frac{\partial \widetilde{\mathrm{MSE}}}{\partial w^{\texttt{ex}}_{S,S}} \mathbf{I}_F \end{bmatrix} \quad . \tag{65}$$

2) Calculation of $\frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{W}^{\texttt{emit}}}$: Let us examine the structure of $\mathbf{W}^{\texttt{emit}}$ in Eq. (49). Here, like in $\mathbf{W}^{\texttt{excite}}$, the actual number of degrees of freedom is less than $SF \times SF$. There are S blocks in the diagonal having $F \times F$ elements each, while other elements have null value. Therefore, the formulation in Eq. (62) is inefficient. Similarly to the case of $\mathbf{W}^{\texttt{excite}}$, we can exploit the sparsity in the structure of $\mathbf{W}^{\texttt{emit}}$: we can differentiate $w_{p,q}$ in Eq. (62) only with respect to the elements $w^{\texttt{em}}_{(s),r,f}$. In this case, we divide $\mathbf{G}$ into blocks of $F \times SF$ sub-matrices $\widetilde{\mathbf{G}}^{\texttt{em}}_{s}$, where s indexes a block of $\mathbf{G}$:

$$\mathbf{G} = \begin{bmatrix} \widetilde{\mathbf{G}}^{\texttt{em}}_{1} \\ \vdots \\ \widetilde{\mathbf{G}}^{\texttt{em}}_{s} \\ \vdots \\ \widetilde{\mathbf{G}}^{\texttt{em}}_{S} \end{bmatrix} \quad . \tag{66}$$

We also divide $\mathbf{W}^{\texttt{excite}}$ into blocks of $F \times SF$ sub-matrices $\widetilde{\mathbf{B}}_s$, where s indexes a block of $\mathbf{W}^{\texttt{excite}}$:

$$\mathbf{W}^{\texttt{excite}} = \begin{bmatrix} \widetilde{\mathbf{B}}_{1} \\ \vdots \\ \widetilde{\mathbf{B}}_{s} \\ \vdots \\ \widetilde{\mathbf{B}}_{S} \end{bmatrix} \quad . \tag{67}$$

Based on Eqs. (50) and (62), we can construct the diagonal blocks of $\frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{W}^{\texttt{emit}}}$:

$$\frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{w}^{\texttt{em}}_{s}} = \widetilde{\mathbf{G}}^{\texttt{em}}_{s} (\widetilde{\mathbf{B}}_s)^T \tag{68}$$

for $s \in [1, \ldots, S]$. Using Eq. (49) the gradient of $\widetilde{\mathrm{MSE}}$ with respect to $\mathbf{W}^{\texttt{emit}}$ is presented as:

$$\frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{W}^{\texttt{emit}}} = \begin{bmatrix} \frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{w}^{\texttt{em}}_{1}} & 0 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \cdots & \vdots \\ \vdots & 0 & \frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{w}^{\texttt{em}}_{s}} & 0 & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 0 & \frac{\partial \widetilde{\mathrm{MSE}}}{\partial \mathbf{w}^{\texttt{em}}_{S}} \end{bmatrix} \quad . \tag{69}$$

VIII. BLEACHING MODEL

Based on Eq. (9), the emitted energy that is captured per pixel for dye d, while excitation source s and emission filter f were active, can be expressed as

$$i^{d}_{s,f} = k^{\text{camera}}_{\text{gray}}\ Q^{\text{camera}}\ \tau^{\texttt{optics}}\ t^{\text{exp}}\ h\ L_s\ \varepsilon_d\ \alpha_{d,s}\ \Phi_d\ e_{d,f}\ c_d \quad , \tag{70}$$

where $e_{d,f}$ $[\%]$ is the integral of the emission spectrum of fluorophore d (peak value is normalized to 1) in the emission filter band f. When multiplexing multiple excitation sources and multiple emission filters, multiple measurements need to be acquired for decoding. There are several limiting processes that can decrease fluorescent emission energy and inhibit the power that is allowed to irradiate the specimen. In this work, we incorporate into the model the phenomenon of photobleaching [6].

A. Bleaching in a single dye, under a single excitation source

Chemical alteration which causes photobleaching decreases the fluorophores' ability to fluoresce. The molecular structure of a specific fluorophore and the experiment conditions affect, in particular, the number of excitation/emission cycles that this fluorophore can undergo before photobleaching.[2] Due to photobleaching, the concentration c_d of the molecules that contribute to the fluorescent emission energy decays in response to the total excitation dosage. This decay is assumed to have exponential nature [7]. Let $P_d^{\texttt{unbleach}}$ be the probability for *not* bleaching a molecule of dye d:

$$P_d^{\texttt{unbleach}} \equiv \frac{c_d(t^{\texttt{total}})}{c_d^0} \quad , \tag{71}$$

where $t^{\texttt{total}}$ is the total accumulated time over all prior measurements, $c_d(t^{\texttt{toatl}})$ is the time dependent expected concentration of fluorescing molecules in the dye, c_d^0 is the initial concentration of the specimen at time $t^{\texttt{total}} = 0$. Therefore,

$$c_d(t^{\texttt{total}}) = c_d^0 \; P_d^{\texttt{unbleach}} \quad . \tag{72}$$

Substituting Eq. (72) into Eq. (70) yields

$$i_{s,f}^{\texttt{bleached}}(t^{\texttt{total}}) = k_{\text{gray}}^{\text{camera}} \; Q^{\text{camera}} \; \tau^{\texttt{optics}} \; t^{\text{exp}} \; h \; L_s \; \varepsilon_d \; \alpha_{d,s} \; \Phi_d \; e_{d,f} \; c_d^0 \; P_d^{\texttt{unbleach}} \quad . \tag{73}$$

According to Eq. (73) the energy that is emitted from the specimen depends on the accumulated exposure time. Therefore, each additional measurement having exposure time $t^{\texttt{exp}}$ increases the total accumulated exposure time by $t^{\texttt{exp}}$. Let p denote the index of measurement number. Up to this measurement, the total accumulated exposure time is

$$t^{\texttt{total}} = (p-1)t^{\texttt{exp}} \quad . \tag{74}$$

---

[2] Another physical limitation is saturation of the fluorophore. Saturation occurs when all the molecules are already in the excited state. Then, an increase in excitation energy is not followed by a corresponding increase in fluorescence emission. This does not impose a practical constraint on the excitation sources. The reason is that in the range of excitation intensities that are used in biological specimen, saturation is not the limiting factor, but photobleaching.

As the total accumulated exposure time increases, the number of molecules that are able to fluoresce is reduced (Eq. 71).

The effective excitation energy falling on the specimen in each measurement depends on the values of $\mathbf{W}^{\mathtt{excite}}$. Thus, from now on, L_s is the maximum laser intensity available for source s. The intensity of source s in each measurement equals $w^{\mathtt{ex}}_{m,s} L_s$. The total accumulated irradiated energy up to measurement p is proportional to the sum of the excitation weights over all previous measurements. Therefore, the total accumulated absorbed energy per molecule of dye d due to all previously active excitation sources until measurement p, can be presented as

$$E_d \equiv \; t^{\mathtt{exp}} \sum_{s=1}^{S} L_s \; \alpha_{d,s} \sum_{p'=1}^{p-1} w^{\mathtt{ex}}_{m,s} \; . \tag{75}$$

Here $k_d^{\mathtt{bleach}}$ $[\% \ \mathrm{J}^{-1}]$ is the percent of molecules that undergo bleaching, per absorbed energy.[3] In Eq. (75) the index m is related to p' by $m = \lceil \ p'/F \rceil$, similarly to Eq. (52). Based on Eqs. (74) and (75) we can express the probability for not bleaching dye d by the time of measurement p as

$$P_d^{\mathtt{unbleach}}(p) = e^{-(p-1) \ E_d \ k_d^{\mathtt{bleach}}} \quad . \tag{76}$$

Using Eqs. (73) and (76) the emitted energy from dye d that is captured in measurement p while source s and filter band f are active (no multiplexing in this particular measurement) is:

$$i_{s,f}^{\mathtt{bleached}}(p) = k_{\mathrm{gray}}^{\mathrm{camera}} \ Q^{\mathrm{camera}} \ \tau^{\mathtt{optics}} \ t^{\mathrm{exp}} \ h \ L_s \ \varepsilon_d \ \alpha_{d,s} \ \Phi_d \ e_{d,f} \ c_d^0 \ P_d^{\mathtt{unbleach}}(p) \quad . \tag{77}$$

B. Bleaching in multiple dyes, under a single excitation source

Let D different dyes be present in the specimen. Therefore, using Eqs. (73) and (13), the energy that is emitted per pixel from the specimen is

$$i_{s,f}^{\mathtt{bleached}}(p) = \kappa^{\mathrm{system}} \ \kappa^{\mathrm{max}} \ L_s^{\mathrm{norm}} \ \sum_{d=1}^{D} \varepsilon_d^{\mathrm{norm}} \ \alpha_{d,s} \ \Phi_d^{\mathrm{norm}} \ e_{d,f} \ c_d^0 \ P_d^{\mathtt{unbleach}}(p) \quad . \tag{78}$$

Using Eqs. (76) and (78), we express the bleaching ratio between $i_{s,f}^{\mathtt{bleached}}(p)$ and $i_{s,f}^{\mathtt{bleached}}(p=1)$:

$$R_{s,f}^{\mathtt{bleach}}(p) = \frac{i_{s,f}^{\mathtt{bleached}}(p)}{i_{s,f}^{\mathtt{bleached}}(p=1)} = \frac{\sum_{d=1}^{D} \varepsilon_d^{\mathrm{norm}} \ \alpha_{d,s} \ e_{d,f} \ \Phi_d^{\mathrm{norm}} \ c_d^0 \ P_d^{\mathtt{unbleach}}(p)}{\sum_{d=1}^{D} \varepsilon_d^{\mathrm{norm}} \ \alpha_{d,s} \ e_{d,f} \ \Phi_d^{\mathrm{norm}} \ c_d^0} \quad , \tag{79}$$

where $i_{s,f}^{\mathtt{bleached}}(p=1)$ is the energy that is captured if no bleaching has occurred. Based on Eq. (79),

$$i_{s,f}^{\mathtt{bleached}}(p) = R_{s,f}^{\mathtt{bleach}}(p) \ i_{s,f}^{\mathtt{bleached}}(p=1) \quad . \tag{80}$$

---

[3] While the energy is low.

C. Bleaching in multiple dyes, and multiplexing

Now, let multiple sources and multiple emission filters be active in each past measurement $p' < p$ (multiplexing). The value that is obtained in each measurement a_p (an element of $\mathbf{a}$), is attenuated as a result of bleaching that had been caused by previous measurements $p' < p$. To express this, we replace the elements of vector $\mathbf{i}$ in Eq. (37) with the elements $i_{s,f}^{\texttt{bleached}}(p)$ from Eq. (80). In this case, each measurement can be expressed as:

$$a_p = \sum_{q=1}^{SF} w_{p,q} i_{s,f}^{\texttt{bleached}}(p) = \sum_{q=1}^{SF} w_{p,q} \; R_{s,f}^{\texttt{bleach}}(p) \; i_{s,f}^{\texttt{bleached}}(p=1) \quad . \tag{81}$$

Here the indexes $\{s, f\}$ are related to the index q by Eqs. (52) and (53). According to Eq. (81), in order to express the attenuation of the energy due to bleaching, we need to multiply each element in the multiplexing matrix $\mathbf{W}$ by the corresponding value of $R_{s,f}^{\texttt{bleach}}(p)$. This would result in the total multiplexing matrix $\mathbf{W}^{\texttt{bleach}}$, accounting for bleaching. Each element in $\mathbf{W}^{\texttt{bleach}}$ is

$$w_{p,q}^{\texttt{bleach}} = w_{p,q} \cdot R_{s,f}^{\texttt{bleach}}(p) \quad , \tag{82}$$

where the indexes $\{s, f\}$ are related to the index q by Eqs. (52) and (53). Thus,

$$\mathbf{a}(x,y) = \mathbf{W}^{\texttt{bleach}} \; \mathbf{i}(x,y) \quad . \tag{83}$$

We wish to decode the energies per wavelength and per emission filter (vector $\mathbf{i}$), as if there was no bleaching. Therefore, the estimation of $\mathbf{i}(x,y)$ is computed by inverting the multiplexing matrix $\mathbf{W}^{\texttt{bleach}}$. From Eq. (83) we get

$$\hat{\mathbf{i}}(x,y) = \left(\mathbf{W}^{\texttt{bleach}}\right)^{-1} \mathbf{a}(x,y) \quad . \tag{84}$$

According to Eq. (79) the values $R_{s,f}^{\texttt{bleach}}(p)$ are calculated using the initial concentrations of dyes c_d^0. The concetrations c_d^0 differ from pixel to pixel. Therefore, the values $R_{s,f}^{\texttt{bleach}}(p)$ are calculated per pixel. Based on Eq. (82), this results in different multiplexing matrices ($\mathbf{W}^{\texttt{bleach}}$) for each pixel. However, it can not be implemented in a real experiment. Therefore, the inital concentrations c_d^0 need to be decoded.

IX. DECODING INITIAL CONCENTRATIONS - UNMIXING

In this section the purpose is to decode the initial concentration c_d^0 for each dye d. Similarly to Eq. (14)

$$\mathbf{i} = \kappa^{\text{system}} \; \kappa^{\max} \; \mathbf{X} \; \mathbf{c}^0 \quad , \tag{85}$$

where $\mathbf{X}$ is the mixing matrix of size $SF \times D$ which has

$$x_{q,d} = L_s^{\text{norm}}\ \varepsilon_d^{\text{norm}}\ \alpha_{d,s}\ \Phi_d^{\text{norm}}\ e_{d,f} \tag{86}$$

as its elements. The indexes $\{s, f\}$ are related to the index q by Eqs. (52) and (53) . The vector $\mathbf{c}^0$ of length D is the vector of the initial concentrations. It has c_d^0 as its elements.

Based on Eqs. (78,85) and (86), the vector of the captured energies up to measurement p, $\mathbf{i}^{\texttt{bleached}}(p)$, can be expressed as

$$\mathbf{i}^{\texttt{bleached}}(p) = \kappa^{\text{system}}\ \kappa^{\max}\ \mathbf{X}^{\texttt{bleach}}(p)\ \mathbf{c}^0 \quad , \tag{87}$$

where $\mathbf{X}^{\texttt{bleach}}(p)$ is a mixing matrix accounting for bleaching. Each element in $\mathbf{X}^{\texttt{bleach}}(p)$ equals

$$x_{q,d}^{\texttt{bleach}}(p) = x_{q,d} \cdot P_d^{\texttt{unbleach}}(p) \quad . \tag{88}$$

We wish to decode the initial concentrations (vector $\mathbf{c}^0$) from a vector of multiplexed measurements $\mathbf{a}$. Using Eqs. (37) and (87) each multiplexed measurement can be expressed as

$$a_p = \mathbf{w}_p\ \mathbf{i}^{\texttt{bleached}}(p) = \kappa^{\text{system}}\ \kappa^{\max}\ \mathbf{w}_p\ \mathbf{X}^{\texttt{bleach}}(p)\ \mathbf{c}^0 \quad , \tag{89}$$

where $\mathbf{w}_p$ is the p's row of the multiplexing matrix $\mathbf{W}$. Therefore, we can define a different multiplexing matrix $\mathbf{W}^{\texttt{bleach}}$ that accounts for bleaching. Each row p in $\mathbf{W}^{\texttt{bleach}}$ is

$$\mathbf{w}_p^{\texttt{bleach}} = \mathbf{w}_p\ \mathbf{X}^{\texttt{bleach}}(p) \quad . \tag{90}$$

Overall, the vector of multiplexed acquired measurements is

$$\mathbf{a} = \kappa^{\text{system}}\ \kappa^{\max}\ \mathbf{W}^{\texttt{bleach}}\ \mathbf{c}^0 \quad . \tag{91}$$

We wish to decode the concentrations per dye (vector $\mathbf{c^0}$). Therefore, based on Eq. (91) and the fact that $\mathbf{W}^{\texttt{bleach}}$ is generally not square, the estimation of $\mathbf{c^0}(x, y)$ is computed by

$$\hat{\mathbf{c}}^0(x, y) = \frac{1}{\kappa^{\text{system}}\ \kappa^{\max}} \left(\mathbf{W}^{\texttt{bleach}^T}\mathbf{W}^{\texttt{bleach}}\right)^{-1} \mathbf{W}^{\texttt{bleach}^T}\mathbf{a}(x, y) \quad . \tag{92}$$

As mentioned before, the estimation of the concentrations up to a constant value is good enough for the unmixing purposes. Therefore, we wish to decode $\tilde{\mathbf{c}} = \hat{\mathbf{c}}\ \kappa^{\text{system}}\ \kappa^{\max}$.

Once the initial concentrations per pixel are available, we can synthesize the values of energies that would have been captured per excitation source and emission filter as if there was no bleaching (Eq. 85). This way we obtain an image which has higher SNR and which is componsated for bleaching. Then, we wish to compare the synthesized results with the ones decoded from multiplexed measurements based on Eq. (38) (no bleaching compensation). If the results of both cases (with bleaching compensation and without) are approximately close, then the conclusion is that bleaching is not a significant factor and that we can neglect this phenomenon.

X. OPTIMAL MULTIPLEXING WITH BLEACHING

In the case of Eq. (17), the $\widetilde{\mathrm{MSE}}$ is as in [8]:

$$\widetilde{\mathrm{MSE}}^{\texttt{bleach}} = \frac{1}{D}\mathrm{trace}\left[(\mathbf{W}^{\texttt{bleach}\,T}\mathbf{W}^{\texttt{bleach}})^{-1}\right] \quad , \tag{93}$$

where $\mathbf{W}^{\texttt{bleach}}$ is defined in Eq. (90). We seek the multiplexing matrix $\mathbf{W}$ that minimizes this MSE similarly to section VII.

A. Calculation of Gradients

To facilitate Eqs. (58) and (59) we differentiate $\widetilde{\mathrm{MSE}}^{\texttt{bleach}}$ with respect to $\mathbf{W}^{\texttt{excite}}$ and with respect to $\mathbf{W}^{\texttt{emit}}$, separately. We define $\mathbf{G}^{\texttt{bleach}}$ as the gradient of $\widetilde{\mathrm{MSE}}^{\texttt{bleach}}$ with respect to $\mathbf{W}^{\texttt{bleach}}$. Then, we use $\mathbf{G}^{\texttt{bleach}}$ to calculate the partial derivatives. Eq (60) becomes:

$$\frac{\partial\widetilde{\mathrm{MSE}}^{\texttt{bleach}}}{\partial\mathbf{W}^{\texttt{bleach}}} \equiv \mathbf{G}^{\texttt{bleach}} = \frac{-2}{D}\mathbf{W}^{\texttt{bleach}}(\mathbf{W}^{\texttt{bleach}T}\mathbf{W}^{\texttt{bleach}})^{-2} \quad . \tag{94}$$

We use the same chain rule as in section VII-B in order to calculate the partial derivatives:

$$\frac{\partial\widetilde{\mathrm{MSE}}^{\texttt{bleach}}}{\partial w_{k,l}^{\texttt{excite}}} = \sum_p\sum_d g_{p,d}^{\texttt{bleach}} \cdot \frac{\partial w_{p,d}^{\texttt{bleach}}}{\partial w_{k,l}^{\texttt{excite}}} \tag{95}$$

$$\frac{\partial\widetilde{\mathrm{MSE}}^{\texttt{bleach}}}{\partial w_{h,j}^{\texttt{emit}}} = \sum_p\sum_d g_{p,d}^{\texttt{bleach}} \cdot \frac{\partial w_{p,d}^{\texttt{bleach}}}{\partial w_{h,j}^{\texttt{emit}}} \quad , \tag{96}$$

where $g_{p,d}^{\texttt{bleach}}$ is an element in $\mathbf{G}^{\texttt{bleach}}$ and $w_{p,d}^{\texttt{bleach}}$ is an element in $\mathbf{W}^{\texttt{bleach}}$. According to Eqs. (88) and (90)

$$w_{p,d}^{\texttt{bleach}} = P_d^{\texttt{unbleach}}(p)\sum_{q=1}^{SF} w_{p,q}\; x_{q,d} \quad . \tag{97}$$

1) Calculation of $\frac{\partial \widetilde{\mathrm{MSE}}^{\texttt{bleach}}}{\partial \mathbf{W}^{\texttt{excite}}}$*:* Similarly to section VII-B.1, we exploit the sparsity and redundancy in the structure of $\mathbf{W}^{\texttt{excite}}$: we can differentiate $w^{\texttt{bleach}}_{p,d}$ in Eq. (95) only with respect to the elements $w^{\texttt{ex}}_{m,s}$. Therefore,

$$\frac{\partial \widetilde{\mathrm{MSE}}^{\texttt{bleach}}}{\partial w^{\texttt{ex}}_{m,s}} = \sum_{p} \sum_{d} g^{\texttt{bleach}}_{p,d} \cdot \frac{\partial w^{\texttt{bleach}}_{p,d}}{\partial w^{\texttt{ex}}_{m,s}} \tag{98}$$

According Eqs. (75) and (76), $P^{\texttt{unbleach}}_{d}(p)$ depends on the sum of the excitation weights ($w^{\texttt{ex}}_{m,s}$) over all previous measurements. Recall the special structure of $\mathbf{W}^{\texttt{excite}}$ where each excitation weight $w^{\texttt{ex}}_{m,s}$, corresponding to excitation source s, is repeated F times for all emission filter bands (Eq. 47). Therefore, $P^{\texttt{unbleach}}_{d}(p)$ is derived with respect to $w^{\texttt{ex}}_{m,s}$. Based on Eqs. (51) and (97):

$$\frac{\partial w^{\texttt{bleach}}_{p,d}}{\partial w^{\texttt{ex}}_{m,s}} = \frac{\partial P^{\texttt{unbleach}}_{d}(p)}{\partial w^{\texttt{ex}}_{m,s}} \sum_{q=1}^{SF} w^{\texttt{em}}_{(m),r,f} w^{\texttt{ex}}_{m,s} x_{q,d} + P^{\texttt{unbleach}}_{d}(p) \sum_{f'=1}^{F} w^{\texttt{em}}_{(m),r,f'} x_{q',d} \quad , \tag{99}$$

where all the indexes are related according to Eqs. (52,53) and (54) and $q' = f + (s-1)F$. Using Eqs. (76) and (75)

$$\frac{\partial P^{\texttt{unbleach}}_{d}(p)}{\partial w^{\texttt{ex}}_{m,s}} = -[p-1] P^{\texttt{unbleach}}_{d}(p) k^{\texttt{unbleach}}_{d} \frac{\partial E_d(p)}{w^{\texttt{ex}}_{m,s}} \quad , \tag{100}$$

where

$$\frac{\partial E_d(p)}{\partial w^{\texttt{ex}}_{m,s}} = \begin{cases} t^{\texttt{exp}} L_s \alpha_{d,s} \left[p - 1 - \left(\left\lceil \frac{p-1}{F} \right\rceil - 1\right) F\right] w^{\texttt{ex}}_{m,s} & p > 1 \\ 0 & p = 1 \end{cases} . \tag{101}$$

The structure of $\frac{\partial \widetilde{\mathrm{MSE}}^{\texttt{bleach}}}{\partial \mathbf{W}^{\texttt{excite}}}$ is similar to the structure of $\mathbf{W}^{\texttt{excite}}$.

2) Calculation of $\frac{\partial \widetilde{\mathrm{MSE}}^{\texttt{bleach}}}{\partial \mathbf{W}^{\texttt{emit}}}$*:* Similarly to the case of $\mathbf{W}^{\texttt{excite}}$, we can exploit the sparsity in the structure of $\mathbf{W}^{\texttt{emit}}$: we can differentiate $w^{\texttt{bleach}}_{p,d}$ in Eq. (96) only with respect to the elements $w^{\texttt{em}}_{(m),r,f}$. Therefore,

$$\frac{\partial \widetilde{\mathrm{MSE}}^{\texttt{bleach}}}{\partial w^{\texttt{em}}_{(m),r,f}} = \sum_{p} \sum_{d} g^{\texttt{bleach}}_{p,d} \cdot \frac{\partial w^{\texttt{bleach}}_{p,d}}{\partial w^{\texttt{em}}_{(m),r,f}} \tag{102}$$

Here there is no need to derive $P^{\texttt{unbleach}}_{d}(p)$ with respect to $w^{\texttt{em}}_{(m),r,f}$. The reason is that unlike in the case of excitation, bleaching is not influenced by the emission filters. Thus, based on Eqs. (51) and (97)

$$\frac{\partial w^{\texttt{bleach}}_{p,d}}{\partial w^{\texttt{em}}_{(m),r,f}} = P^{\texttt{unbleach}}_{d}(p) \sum_{s=1}^{S} w^{\texttt{ex}}_{m,s} x_{q',d} \quad , \tag{103}$$

where all the indexes are related according to Eqs. (52,53) and (54) and $q' = f + (s-1)F$. The structure of $\frac{\partial \widetilde{\mathrm{MSE}}^{\texttt{bleach}}}{\partial \mathbf{W}^{\texttt{emit}}}$ is similar to the structure of $\mathbf{W}^{\texttt{emit}}$.

XI. CALIBRATION OF $P_d^{\texttt{unbleach}}$

Before performing the experiment using Eq. (91), the elements of the matrix $\mathbf{W}^{\texttt{bleach}}$ must be calibrated for each dye independently. According to Eqs. (88) and (90) the only unknown parameter, which is dependent on p, is $P_d^{\texttt{unbleach}}(p)$. Therefore, we need to calibrate it, in order to obtain the elements of the matrix $\mathbf{W}^{\texttt{bleach}}$.

From Eqs. (74,76) and (77), the energy acquired at measurement p for source s, dye d and filter f is

$$i_{s,f}^{\texttt{bleached}}(p) = k_{\text{gray}}^{\text{camera}}\ Q^{\text{camera}}\ \tau^{\texttt{optics}}\ t^{\text{exp}}\ h\ L_s\ \varepsilon_d\ \alpha_{d,s}\ \Phi_d\ e_{d,f}\ c_d^0\ e^{-(p-1)\ E_d\ k_d^{\texttt{bleach}}} \quad . \tag{104}$$

Let us write Eq. (104) in a compact way combining the various parameters

$$i_{s,f}^{\texttt{bleached}}(p) = be^{-(p-1)\beta} \quad , \tag{105}$$

where

$$b = k_{\text{gray}}^{\text{camera}}\ Q^{\text{camera}}\ \tau^{\texttt{optics}}\ t^{\text{exp}}\ h\ L_s\ \varepsilon_d\ \alpha_{d,s}\ \Phi_d\ e_{d,f}\ c_d^0 \tag{106}$$

and

$$\beta = E_d k_d^{\texttt{bleach}} \quad . \tag{107}$$

Our purpose is to perform a calibration experiment to find β and then to extract $k_d^{\texttt{bleach}}$. Taking $\ln$ of both sides of Eq. (105) yields the following equation for a line, having β as a slope:

$$\ln[i_{s,f}^{\texttt{bleached}}(p)] = -(p-1)\beta + \ln(b) \quad , \tag{108}$$

For each dye, we perform a calibration experiment. We measure $i_{s,f}^{\texttt{bleached}}(p)$ as a function of measurement number p while maintaining the same source intensity for all measurements. Then, we plot the $\ln$ of acquired measurements $i_{s,f}^{\texttt{bleached}}(p)$ as a function of measurement number p and compute its slop to obtain β. The value of $k_d^{\texttt{bleach}}$ is then computed based on Eq. (107), where E_d is known from Eq. (75).

REFERENCES

[1] Y. Y. Schechner, S. K. Nayar and P. N. Belhumeur, 2003, “A Theory of Multiplexed Illumination”,In Proc. IEEE ICCV, Vol. 2, 808–815.

[2] N. Ratner and Y. Y. Schechner, 2007, “Illumination multiplexing within fundamental limits”, In Proc. IEEE CVPR.

[3] H. V. Henderson and S. R Searle, 1981, “The vec-permutation matrix, the vec-operator and Kronecker products: a review”, Linear and Multiliniar Algebra, 9, pp. 271-288.

[4] D. P. Bertsekas, 1999, “Nonlinear Programming”, Athena Scientific.

[5] K. B. Petersen and M. S. Pedersen, 2008, “The Matrix Cookbook”,Technical University of Denmark.

[6] R. A. Mathies and L. Stryer, 1986, “Single-molecule fluorescence detection: a feasibility study using phycoerythrin” Chapter 6 in Applications of fluorescence in the biomedical sciences D. L. Taylor, Editor. Alan R. Liss, Inc, New York, NY pp. 129-140.

[7] C. Eggeling, J. Widengren, R. Rigler and C. A. M. Seidel, 1998, “Photobleaching of Fluorescent Dyes under Conditions Used for Single-Molecule Detection: Evidence of Two-Step Photolysis”, Analytical Chemistry, Vol. 70, No. 13, July 1.

[8] M. Harwit and N. J. A. Sloane, 1979, “Hadamard Transform Optics”, Academic, New York.