

Article

Not peer-reviewed version

Extracting the Most Relevant Information from Bio-Photonic Data: A Constrained Maximum Entropy Methodology

[Rosa Bernardini Papalia](#) *

Posted Date: 27 November 2024

doi: 10.20944/preprints202411.2128.v1

Keywords:

bio-photons emission; quantum data; constrained maximum entropy



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Extracting the Most Relevant Information from Bio-Photonic Data: A Constrained Maximum Entropy Methodology

Rosa Bernardini Papalia

Department of Statistical Sciences, University of Bologna, Italy; rossella.bernardini@unibo.it

Abstract: This paper introduces a Constrained Maximum Entropy (CME) approach to extract knowledge from bio-photonic data. More specifically, we discuss the main issues related to this new type of data and show the potential of the CME methodology to introduce both *a priori* knowledge and data constraints to efficiently analyze bio-photon data. The advantage in our view is that it is possible to find the most unbiased bio photons distribution, that is the distribution which has the maximum entropy among the class of distributions satisfying constraints imposed by given knowledge and being least uncommitted to information not yet available. Then, we go a step further and suggest that the quantitative and qualitative constraints of the CME formulation based on accurate bio-photon emissions will provide a powerful new tool to monitor changes in biological systems and will be used to identify potential unstable states and evaluate the effect of a new treatment on a biological system.

Keywords: bio-photons emission; quantum data; constrained maximum entropy

1. Introduction

Despite the availability of advanced data acquisition technology, extracting the most relevant information from bio-photonic data for its efficient use in different fields remains a non-trivial task. The massive amount of data that is often difficult and time-consuming to interpret represent a big challenge. It is an instinct to confine information to that only which can be seen with the human eye. Nevertheless, bio photons provide useful information that can only be obtained from a deeper quantitative analysis of the same data. Scattering of light in biological systems contributes to what is called polarized light. Taking measurements of the polarized nature of light building a model and simulating the propagation of light can yield information that is far too complex to interpret qualitatively. The development of inverse algorithms and models to extract useful information may represent a potential tool. An inverse problem is one in which one seeks to determine the cause given an effect. In the context of this research, it is sought to determine the optical properties of a given biological sample and how it may be related to some condition or disease. This would involve making comparisons with measured data and some form of minimizing the error between the model and the actual data. The data and the optical properties that caused it can usually be quite complex and thus inverse algorithms must be formulated with this in mind. Many problems in bio-photonics will involve ill-defined matrices and may be non-linear. Step one in solving an inverse problem must always be to simplify the problem to something that is solvable and useful. With proper extraction of relevant data, the result can be a powerful tool for general decision guidance and for complex decisions that use as the deciding factor non-invasive light-based techniques.

On its power of visualization, measurement and quantification at micro and nano scales, bio-photonics has contributed significantly to the discoveries and understanding in life sciences. Inevitably, the complexity and amount of data to be generated from various bio-photonic imaging and sensing techniques have grown by leaps and bounds. Thus, the potential to extract useful information from the data and transform it into knowledge has increased the data analysis importance in bio-photonics. It means that data shall not be interpreted in the light of the method used in its acquisition (hypothesis and expected outcome) to avoid information biases. Instead, the

data shall be used to derive the best possible answer to the initial questions and to provide a solid foundation for the growth of knowledge with minimum assumptions. Only through effective and accurate analysis, bio-photonic data will turn into trusted information, providing the basis for new knowledge and potentially leading to the big discoveries. Data analysis importance in bio-photonics can be seen through its direct relatability to human health and wellbeing of a general biological system (animal, plants...). Finally, the trusted information derived from data will greatly influence the decision and policy makers.

The objectives of this work are: (i) to provide a method for the identification of the most crucial problems in a bio-photonic measurements. The focus will be the development of methods to "rank" data points in terms of their importance to the desired information; (ii) to develop tools through an Information Theoretic based framework that will assist a bio-photonic researcher in getting the most out of their experiments and data analysis.

2. Background of Bio-Photonic Data

Real bio-photonic data sets are derived from the analysis of light propagation in biological systems. The data sets are typically very large (128x128 to 512x512 and larger) and currently are analyzed point by point and only provide information on scattering. This leads to the first and most important challenge in analyzing bio-photonic data, that of developing an understanding of the various contributions to light transport in biological systems. Light transport in tissues is a complex process involving absorption, scattering, fluorescence, and luminescence. The development of a comprehensive model of light transport in tissues and how it can be best characterized by specific types of measurements is an open problem. The lack of a comprehensive model of what data means in a biological sense is another challenge of bio-photonic data analysis (Luo, 2020; Pan et al., 2021; Ahi et al., 2020; Rodrigues & Hemmer, 2022; Yoon et al., 2022; Chen et al., 2022; Balasundaram et al., 2021; Fanous et al., 2024; Verma & Gupta, 2023).

Most biophotonic measurements are taken without a direct corresponding anatomical location. Without the ability to locate where a measurement was taken, it is difficult to properly interpret the measured data in a biological sense and relate it to other measurements or known physiological conditions. A different type of challenge which is specific to bio-photonic data analysis is its interdisciplinary nature. Most bio-photonic measurements are taken by physicists or engineers who may not have a strong biological background, or if they are biologists or medical researchers the data may not be well characterized from a physical standpoint. This type of data is becoming more prevalent with the increasing popularity of purely optical clinical measurements and is typical of research data which often blurs traditional disciplinary boundaries. Interdisciplinary research has its own inherent challenges of communication between researchers of different backgrounds and effectively translating and understanding data which is not in one's native research domain. This can cause difficulties when bio-photonic data is to be analyzed or interpreted by someone who does not have a strong understanding of what the data actually means in a biological sense.

The bio-photonic data is information obtained and stored in a digital form from a biological sample using an optical method. To expand on this general definition, one can consider many types ranging from microscopic and spectroscopic images and spectra, to data obtained from a light based therapeutic intervention. The following are key characteristics of bio-photonic data. The way it is stored and manipulated means that it is often easy to exchange and also may be analysed on different sites to those of data acquisition.

Bio-photonic data is usually non-invasive, digital optical methods can be used to monitor the effects of treatments in longitudinal studies and there is potential to apply new methods immediately to biological systems. Because bio-photonic data is obtained using optical means it is often quite visual and can be easily interpreted by those who are not specialists in the data's area of research.

This is important for effective communication between researchers who are studying different topics but can share knowledge and methods. Finally and most importantly, there is a tremendous potential for the improvement of treatment of unstable states using methods that have analyzed bio-photonic data.

2.1. Characteristics of Bio-Photonic Data

Biophotonic data can encompass information from a variety of sources in its attempt to understand the interactions of life and light at a fundamental level. These sources include but are not limited to data obtained from imaging the inner structure of materials and from the way light scatters when it interacts with materials. This data can be obtained via several means including videos, still images, digital data, or analog data. This may include the measurement of signaling events in a biological system using light. Photonics is the science and technology of generating, controlling, and detecting photons, which are particles of light. There is a great deal of overlap between photonics and light-based technology, and biophotonics is often used to denote the application of optical science to biological systems.

Bio-photonic data is obtained by cameras, from electronic detectors that measure the number of photons (resolution elements) at a point on the sample, or from spectrophotometric measurements that record the number of photons of a specific wavelength that are absorbed or elastically scattered at a particular point. Some factors to be considered in the choice of source include the information content and the range of specificity of the measurement, the depth or area from which information is desired, and the possible effects on alterations to the sample.

The primary information content in an image is the distribution of intensity of the detected photons as a function of the position on the sample. This may be used to form a visual image, or it may be recorded and used as data to determine certain characteristics of the sample, such as its absorption or scattering properties. The type of image obtained is determined by the field and detector that are used with the source, and a variety of techniques are available. For example, an image in which the intensity at each point represents the color of the sample, or a monochromatic image of the scattering of light at a specific wavelength, or an image in which the color and intensity represent the spectra of emitted or reflected light at each point can be obtained.

While spectral parameters are an important step forward in the characterization of a photon, the most powerful part of bio-photonic data analysis is the derivation of "what you can't measure". This is done through derived parameters, which can be accomplished by formulating a specific question and then developing a model to answer that question. This model should incorporate some a priori knowledge about the system in the form of known physics and make some predictions.

Spectral parameters relate to the fact that photons can have different energies. The wavelength of a photon is inversely proportional to its energy, i.e., $E=hc$, where E is energy, h is Planck's constant, and c is the speed of light. Photons can be either "bunched" together or emitted singly, and according to the "bunching" theory, high-energy photons are emitted at an earlier time than lower-energy photons. Therefore, it can be seen that it is possible to have spectral resolution in a temporal data set. High spectral resolution is useful in differentiating between different types of tissue due to the fact that some tissues have different absorption and scattering coefficients for different wavelengths of light. This may lead to the generation of an elastic scattering spectrum, which could give information about the depth of light penetration into the tissue. A common example of a bio-photonic data set with spectral resolution is the reflectance or fluorescence spectrum from a specific type of tissue.

The spatial and temporal parameters are the most straightforward parameters to understand and can be attributed to the location and timing of the data-taking process. Temporal information is also important because some photons are measured over a period of time, i.e., fluorescence lifetime measurements.

2.2. Pre-Processing Techniques for Bio-Photonic Data

Pre-processing routines are vital before the data can be modeled, in the sense that it reduces noise and produces reliable input for further processing. Noise is any data that is uninterpretable, where in bio-photonic imaging it is usually caused by low light images and auto-fluorescence from the sample (Hibbs, D, 2004). Noise usually arises from bio-photonic data due to variations in the data capture process. Implementing imaging techniques can be used to significantly reduce these certain types of noise, but may result in an increased presence of other noise types. The objective of signal processing is to extract the signal which is of interest by defining the signal and noise. Image will

degrade either by additive or multiplicative noise, and so the aim is to clean noise from the image without destroying the original signal. Signal filtering is a technique used to modulate the information content of a signal. An ideal filtering technique will eliminate the noise and leave the original signal intact. There are two main types of filtering: linear and non-linear. In linear filtering, the information content of the signal is modified. This may make the interpretation of the information content more complex.

Noise removal techniques are generally divided into two categories: data-driven and hard-modelling methods. Data-driven noise removal methods place primary focus on extracting information from the data about the nature of the noise and use this information to modify the data to a cleaner form. This is often achieved by making few assumptions about the data and using known statistical and mathematical theory to estimate the true data. If enough is known about the nature of the noise, it may be possible to develop a method to remove it by constructing a model specific to the noise type. It is possible to assess the effectiveness of the noise removal approach using statistical analysis and by comparing the noise-free data to the original data, an error assessment can be made.

A low-pass filter is a typical filter used in noise removal, which passes through all components of the image that have a frequency lower than a certain cut-off frequency and attenuates components with frequencies higher than the cut-off frequency (Jourlin et al. 2012). Classical image thresholding is a well-known method of image segmentation which can be used to distinguish objects from a background. However, this method is not well-suited to images that contain inhomogeneous objects. In such instances, implementation of the maximum entropy method is preferable. Temporal noise, and some forms of random noise can be reduced by using wavelet transform. Wavelet transform performs a signal decomposition into a set of basis functions. Each basis function has a different level of localization in time and frequency. Wavelet thresholding has an advantage over Fourier analysis since data and functions can be analyzed on more than one scale. By applying a threshold to the wavelet coefficients, the noisy coefficients can be removed and the signal can be reconstructed with reduced noise. Finally, Maximum Entropy thresholding seems to be an efficient method for segmentation of fluorescence lifetime images, and has the advantage that it is based on a consistent statistic. Assuming that the data represents a single experimental condition and is obtained from multiple scans, a systematic method to reduce the error in the data is to normalize the data. The goal of data normalization is to remove the variation due to the measurement conditions from the data in order to compare the signal to a constant baseline. The simplest data normalization technique is to divide the signal by its reference value to find the intensity as a multiple of the reference signal. Although simple, this method is, in general, too restrictive as it assumes that the noise on the signal and the baseline are equal. In reality, this may not be the case and the real desire is to compare the signal to a noise-free baseline, which does not have to have the same shape as the signal baseline. It does, however, have the advantage of allowing comparison of images obtained under different conditions and is also reversible, so the transformed image can be compared to the original data.

Traditional fitting methods such as robust fitting, component model separation, wavelet analysis, and singular value decomposition perform acceptably in areas where model simplification may adequately describe some observed behaviors. However, they all lack utilization of learned information, which can be encoded in the form of Shannon entropy-based analytical probability density models. Consequently, they often suffer in predicting previously unseen behaviors, and there exist validity problems in the analysis of multiple modes or colors of data. Semiclassical and intrinsic Bayes classifiers perform well in unsupervised and supervised classification of observations when Gaussian assumptions are satisfied.

3. Constrained Maximum Entropy Methods in Biophotonics Data Analysis

Along with the development of experimental biophotonics methods, the urgent need for efficient tools and techniques to analyze acquired data coexists. The underlying reasons are reflected in the increasing complexity of data obtained via novel experimental contributions to the field and limitations in current bioimaging analyses. Involvement of novel and innovative tools and

methodologies from other scientific fields in biological investigations appears as the modern trend in analytical and statistical biophotonics.

Photonic methods allow the development of noninvasive examination tools that offer high precision measurements and can be used to monitor and analyze biological systems.

The main contents of this section are a concise overview of the most important issues of maximum entropy methods, a list of the types of problems to which these methods have been applied satisfactorily.

GME methods grow out of the development, in the 1950s, of the principle of maximum caliber. According to this principle, given primary knowledge expressed as expected values for a collection of macroscopic properties of a physical system, the probability distribution—prime distribution for its possible microstates—partitioning the states into properties that are compatible with values that match those expected and those that do not—is the one with the greatest Shannon entropy (Shannon, 1948).

This result complements other fundamental results of statistical mechanics, but it also generalizes them greatly. While there exist efficient measurements for a limited subset of the properties describable in principle and for rare states of the system, there is potential information in the remaining possible microstates; that potential defines alternative, stable, and in general qualitatively good probability distributions, sometimes very different from the equilibrium one. Making use of that information is thus a good enough method for developing novel testable hypotheses, for checking statistical models and evaluating their confidence levels while avoiding undue bias. In addition, mutual information is an especially appealing association measure for high-throughput studies of intracellular networks due to its information-theoretic characteristics, more global nature, and ability to tease out nonlinear relationships in the presence of noise. It has been previously used in experimental and computational studies of diverse systems, from simple preservation to more complicated ones where traditional methods might be misled by the complex issue of multiple hypotheses, particularly if experimental control samples are unavailable, the multiplicity of comparisons is circumscribed, and the high dimensionality of the dataset is responsible for strongly skewed feature distributions. Crucial problems for a successful application of the mutual information measure are rapid and precise association estimation with acceptable false association control and adequate noise robustness of the association detection. It has been revealed in numerous communication and information-theoretic function approximations and limit theorems that the entropy functional (and the information functional for the conditional case) have Gaussian self-averaging averages in the large sample size limit. With the increase of sample size, the dependence of the associated estimation on specific estimation algorithms tends to disappear.

3.1. Overview of Constrained Entropy Methods

Maximum entropy modeling can be employed in cases where imperfect information is available to both select natural features and also to ensure an unambiguous posterior prediction (Golan et al., 2023; Golan, 2018; Golan, 2008; Golan et al., 1996; Bernardini Papalia et al., 2021; Bernardini Papalia et al., 2018). The theory of maximum entropy learning assumes a set of mappings from observed events to constraint features.

The problem of learning an unknown probability function p from a set of observed data $X_1, \dots, X_n$ is of fundamental importance in statistics. We consider a relatively complex feature space and a highly non-parametric model.

Although prior knowledge regarding X , which we group as features $f_1, \dots, f_j$, may be available, the only information allowed in the GME model formulation is a list of constraints. Such information, if accurate, is clearly useful for learning a better posterior. Even when the observation space X is simple, effectively using such potentially relevant information without mistakenly allowing the hypothesis class to overfit is a more advanced problem.

Entropy is a measure of the uncertainty that a random variable in a particular state exists. For a continuous probability distribution, the entropy is defined by a differential equation that has unique properties for probability distributions that are equilibrium solutions for maximization of entropy

under appropriate constraints. The maximum entropy principle gives a consistent and objective approach to the construction of probability distributions based only on partial information and basic principles of probability. Maximum entropy utilizes all of the provided limited information and, in general, leads to more robust probability and invariant density distributions than any other methods.

Let Z be a random variable that lies in a certain state z with a probability of $p(z)$ where $Z \in [z_{\min}, z_{\max}]$. It can be assumed that an approximative probability density function estimated from a given single data set can describe such a random variable Z by methods like histograms, kernel density estimators, or maximum likelihood estimation. Although a derived $p(z)$ will ensure some reduced entropy, note that for a model that is based on some data and an underlying probability, the model is completely determined and unique. All other models depend on an appropriate modification of its probability, which is the main result of this work. The entropy $H[p(z)]$ of Z , as implied by $p(z)$, can be formulated by using the definition of Shannon's entropy with a normalized constant N , i.e., $H[p(z)] = -N \sum p(z) \ln(p(z))$, where R is given as $R = 2$ or $R = 10$ if the entropy is assumed to be in log using 2 or 10 as the base, respectively.

3.2. Applications in Data Analysis

One of the major challenges in the field of biophotonics is the systematic data analysis and the extraction of knowledge from the generated data. To overcome this obstacle in the field, general maximum entropy methods, often in the form of negative Laplace transform or different types of inversion methods, is here proposed for data obtained by light transmission at a given set of spatial sampling locations.

Furthermore, general maximum entropy methods, often formulated in the form of a weak-constraint minimization problem, can be utilized for space-time hyperspectral imaging of biological systems, and for utilizing a new class of nodes in conjunction with compressive sampling to gather significantly more information about the labels from the same number of measured diffraction patterns than can be gathered with conventional methods. That is, more successful than what can be achieved via compression sampling alone.

The constrained maximum entropy method is an approach to model problems where the probabilities are known only on a countable subset of the sample space. We replace the unknown information on the complement of the countable subset by assigning estimated probabilities derived from a regularized entropy. The counting probabilities act as a horizon toward the past, and there are the equilibrium of the maximum Boltzmann-Gibbs entropy. These estimated probabilities are derived from maximization entropy within a multifractal modeling. The generalized maximum entropy uses the principle of maximum entropy to solve probability models where the probabilities are unknown on a countable set. To construct a maximum entropy probability model on a countable state space, an empirical probability measure calculated on the experimental data is used, and a class of probability measures is defined. Based on the empirical probability measure, it defines an M -probability model, which is mainly concentrated on the empirical probability measure and maximizes its entropy. Differing from other methods of estimation and inference, which seek an estimative distribution based on a maximum likelihood viewpoint, the maximum entropy method looks for the distribution whose divergence to the uniform distribution is minimal, among all distributions that satisfy whatever partial information is available in the form of expectations.

The main advantage of the constrained entropy method is its ability to deliver distributions and give a suitable meaning to the system of constraints based on incomplete information. The maximum entropy method also has some drawbacks associated with its particular form of information. A priori information can be considered.

Analytical constrained entropy methods are very successful in many challenging data interpretation problems where prior knowledge is known a priori and fundamental constraints are available. The application of prior information in the form of a Shannon entropy-based functional combined with a minimum information loss principle makes maximum entropy techniques a potentially more powerful method. The traditional linear and polynomial fitted models describe data with low complexity, but such models are known to singularly fail in high dimensions, skewed data

distribution, or in cases when there is a lack of prior understanding about the subject data. In extremely noisy experiments where the acquisition time is limited and when the signals are partial or truncated, the validity of traditional fitting is questionable.

CME methods are well accommodated to these intrinsic fuzziness and associated estimation variability of biophotonic data and they seem to accommodate the quantum uncertainty present in the collection of the biophotonics image data in order to fully statistically describe the data modeling results and their estimation uncertainty.

The readiness and the benefits of this additional methodological support provide a supplement to the deep dictionary learning pursuit of advanced experiments and reduce the impedance mismatch between the data and the advanced analysis tools.

This method allows for a statistical approach to the restoration of quantum biological system images or pure quantum data within a matrix algebra formalism and has great potential for the research field in many real-life tasks in areas such as quantum biology, quantum information science, and biophotonics.

According to available information the Constrained Maximum Entropy (CME) method models the photon emission of a biological system as a frequency distribution over states and proceeds by ordering the frequency distributions that satisfy the constraints (the information used) by their Shannon information entropy or, when available information suggests a non-uniform prior, by their relative entropy (the Kulback-Leibler divergence). With this information the resulting maximum posterior probability distribution is the frequency distribution that satisfies the constraints that has the highest Shannon information entropy or the minimal Kulback-Leibler divergence and is least uncommitted to information not yet available. In addition, following Golan (2018) and Bernardini Fernandez (2021) this framework can be generalized to allow for noise in the constraints given the uncertainty about the process under study and the nature of count of the variable of interest.

In our context a noise component ε_i for each count observation c_i is included as a constraint in the CME formulation:

$$c_i = p_i + \varepsilon_i \quad (1)$$

In such a case, we assume that the c_i elements are given from two sources: a signal plus a noise term (ε_i) that refers to our uncertainty about the target variable.

Each count c_i is assumed as a discrete random variable that can take M different values. Defining a supporting vector (for the sake of simplicity assumed as common for all the y_i) $z' = [z_1, z_2, \dots, z_M]$ that contains the M possible realizations of the targets with unknown probabilities $p'_{ij} = [p_{i1}, p_{i2}, \dots, p_{iM}]$, c_i can be written as:

$$c_i = \sum_{m=1}^M p_{im} z_m \quad (2)$$

The idea can be generalized in order to include an error term ε_i and define each c_i as:

$$c_i = \sum_{m=1}^M p_{im} z_m + \varepsilon_i \quad (3)$$

We represent uncertainty about the realizations of the errors treating each element ε_{ij} as a discrete random variable with $L \geq 2$ possible outcomes contained in a convex set $\mathbf{v}' = \{v_1, \dots, v_L\}$, which for the sake of simplicity will be assumed as common for all the ε_{ij} . We also assume that these possible realizations are symmetric around zero ($-v_1 = v_L$). The traditional way of fixing the upper and lower limits of this set is to apply the three-sigma rule (see Pukelsheim, 1994). Under these conditions, each ε_i can be defined as:

$$\varepsilon_i = \sum_{l=1}^L w_{il} v_l; \forall i = 1, \dots, T; \quad (4)$$

where w_{il} is the unknown probability of the outcome v_l for the count i .

The model can thus be written in the following terms (5):

$$c_i = \sum_{m=1}^M p_{im} z_m + \sum_{l=1}^L w_{il} v_l \quad (5)$$

The solution to the estimation problem is given by the minimization of the Kullback-Leibler divergence between the posteriors distributions $\mathbf{p}'$ s and the *a priori* probabilities $\mathbf{q}'_i = [q_{i1}, q_{i2}, \dots, q_{iM}]$.

The solution to the estimation problems is given by minimizing the KL divergence between the $\mathbf{p}'$ s and the $\mathbf{q}'$ s. Specifically, the constrained minimization problem can be written as:

$$\underset{\mathbf{p}, \mathbf{W}}{\text{Min}} D(\mathbf{p}, \mathbf{W} \| \mathbf{q}, \mathbf{W}^0) = \sum_{m=1}^M \sum_{i=1}^T p_{im} \ln \left(\frac{p_{im}}{q_{im}} \right) + \sum_{l=1}^L \sum_{i=1}^T w_{il} \ln \left(\frac{w_{il}}{w_{il}^0} \right) \quad (6)$$

Subject to:

$$c_i = \sum_{m=1}^M p_{im} z_m + \sum_{l=1}^L w_{il} v_l; \forall i = 1, \dots, T \quad (7)$$

$$\sum_{m=1}^M p_{im} = \sum_{l=1}^L w_{il} = 1; \quad 1 \forall i = 1, \dots, T; \quad (8)$$

Restrictions (8) are just normalization constraints, whereas (7) reflects the observable information that we have on the count c_i . If we do not have an informative prior, the *a priori* distributions are specified as uniform ($q_i = \frac{1}{M}; \forall m = 1, \dots, M$), which leads to the GME solution. The uniform distribution is usually set as the natural prior $\mathbf{W}^0$ for the error terms.

Following Golan et al. (1996), it is possible to introduce other data constraints in the CME formulation if additional information is available.

Once the respective supporting vectors and the *a priori* probability distributions are set, the estimation can be made in the terms of the following program formulation:

Both for the parameters and the errors, the supporting vectors usually contain values symmetrically centered on zero. If all the *a priori* distributions ($\mathbf{q}^a, \mathbf{q}^b, \mathbf{W}^0$) are specified as uniform, then the GCE solution reduces to the GME one.

To recover the probability vectors $\mathbf{p}_i$, the Lagrangian Function following the matricial form would be:

$$\begin{aligned}
L = & \sum_{m=1}^M \sum_{i=1}^T p_{im} \ln \left(\frac{p_{im}}{q_{im}} \right) + \sum_{l=1}^L \sum_{i=1}^T w_{il} \ln \left(\frac{w_{il}}{w_{il}^0} \right) \\
& + \lambda \left[c - \sum_{m=1}^M p_{im} z_m + \sum_{m=1}^M w_{il} v_l + \right] + \vartheta_i \left[1 - \sum_{m=1}^M p_{im} \right] \\
& + \mu_i \left[1 - \sum_{l=1}^L w_{il} \right]
\end{aligned} \quad (9)$$

with the first order conditions:

$$\begin{aligned}
\frac{\partial L}{\partial p_{im}} = & \sum_{i=1}^T \left(\ln \left(\frac{p_{im}}{q_{im}} \right) + 1 \right) - \lambda \sum_{m=1}^M z_m - \sum_{i=1}^T \sum_{m=1}^M \vartheta_i = 0 \\
& i = 1, \dots, T; m = 1, \dots, M
\end{aligned} \quad (10)$$

$$\begin{aligned}
\frac{\partial L}{\partial w_{il}} = & \frac{\partial L}{\partial w_{il}} \sum_{i=1}^T \left(\ln \left(\frac{w_{il}}{w_{il}^0} \right) + 1 \right) + \lambda \sum_{l=1}^L v_l - \sum_{i=1}^T \sum_{l=1}^L \mu_i = 0 \\
& i = 1, \dots, T; l = 1, \dots, L
\end{aligned} \quad (11)$$

$$\frac{\partial L}{\partial \lambda} = c - \sum_{m=1}^M p_{im} z_m + \sum_{l=1}^L w_{il} v_l = 0 \quad (12)$$

$$\frac{\partial L}{\partial \vartheta_i} = 1 - \sum_{m=1}^M p_{im} = 0; i = 1, \dots, T \quad (13)$$

$$\frac{\partial L}{\partial \mu_i} = 1 - \sum_{l=1}^L w_{il} = 0; l = 1, \dots, L \quad (14)$$

The solution to this system of equations and parameters yields the following solution:

$$\hat{p}_{im} = \frac{q_{im} \exp[\lambda w_{il}]}{\Omega(\hat{\lambda})} = \frac{q_{im} \exp[\hat{\lambda} w_{il}]}{\sum_{m=1}^M q_{im} \exp[\hat{\lambda} w_{il}]} \quad (15)$$

where $\Omega(\hat{\lambda}) = \sum_{m=1}^M q_{im} \exp[\hat{\lambda} w_{il}]$ is a normalization factor and $\hat{\lambda}$ is the estimate of the Lagrange multiplier associated to constraints. The constrained optimization problem can be formulated in

terms of the -unconstrained- dual $L(\lambda) = \lambda Y - \ln \Omega(\lambda)$, depending only on the parameter λ . Using the estimated distribution of previous temporal measurements as prior distributions, it is possible to empirically model the learning that occurs from repeated samples (Bernardini Papalia, 2024).

4. Conclusion and Perspectives

In this work, we have proposed an IT based method to extract relevant information from bio-photonic data. A photonic data is a map of intensity and of how much light is scattered from a sample. There is a wealth of data produced particularly from scattering of light using a variety of different experimental setups. Such data is extremely information rich and has the potential to provide new understanding of nature and balanced state of biological systems and how they are altered or changed. The goal is to develop automated methods to interpret such new data.

There is vast potential for future research in this area due to the many types of light scattering experiments and the various possible biological composition and structural changes in biological systems (Bernardini Papalia et al. 2023). The key findings demonstrate that methods involving a forward computational model can be used to interpret photonic data in terms of specific structural or biochemical changes. This is essential if the full potential of such data is to be realized in addressing complex questions applied to a wide range of experiments. There is potential to develop these methods into a practical tool for applied researchers. This would involve the development of inverse algorithms to estimate pattern difference maps from real data and validation of the methods against experimental studies on biological systems.

Scattering patterns from well-characterized changes can be usefully quantified relative to normal states using pattern difference maps. These are maps of the difference in scattering intensity at each spatial point between unbalanced or changed states and normal/stable state. It is possible to predict such pattern difference maps computationally using an approximation of light transport with a good degree of accuracy. This is important because it provides a forward model for interpreting pattern difference maps. Finally, pattern difference maps can be resolved into images of change in a way which directly represent changes in biological structure. Images of change can potentially provide an entirely new level of understanding of structural changes in bio systems and how they are related to their unbalanced states. Our method can be applied to data from experiments where there is well-established knowledge of the nature of structural or biochemical change.

Within this novel method for extracting the most salient information from bio-photonic data the most significant changes in the data are extracted, and aggregated information about these changes can be obtained using traditional statistical analysis.

The method has wide applicability to many problems in which changes in complex multivariate data need to be detected and interpreted, and thus represents a general advance in the analysis of bio-photonic data. In particular, this method can be effectively utilized in the analysis of spectral data obtained from optical measurements on biological systems.

Future research in data mining from optical measurement systems has many potential directions. One approach that has great potential is the use of data fusion from multiple data sources. For example, many of the biological systems samples that are measured optically also are measured with other instruments related to chemical properties etc.... Fusion of the data from these devices with the optical measurement data has the potential to greatly increase the information that can be obtained from the original sample. In order to facilitate data fusion, it will be important to develop measurement standards in bio-photonics so that data from different instruments and different labs can be meaningfully compared and combined. This is an issue in which interdisciplinary teams must be involved. Fusion of data and estimation by entropy maximization has the potential to engage all types of information even at an aggregate level.

References

Luo, Q., 2020. A brief introduction to biophotonic techniques and methods. Science China Life Sciences. springer.com

- Pan, T., Lu, D., Xin, H., & Li, B., 2021. Biophotonic probes for bio-detection and imaging. *Light: Science & Applications*. nature.com
- Ahi, K., Jessurun, N., Hosseini, M.P. and Asadizanjani, N., 2020. Survey of terahertz photonics and biophotonics. *Optical Engineering*, 59(6), pp.061629-061629. spiedigitallibrary.org
- Rodrigues, E. M. & Hemmer, E., 2022. Trends in hyperspectral imaging: from environmental and health sensing to structure-property and nano-bio interaction studies. *Analytical and Bioanalytical Chemistry*. [HTML]
- Yoon, S., Cheon, S.Y., Park, S., Lee, D., Lee, Y., Han, S., Kim, M. and Koo, H., 2022. Recent advances in optical imaging through deep tissue: imaging probes and techniques. *Biomaterials Research*, 26(1), p.57. science.org
- Chen, X., Lindley-Hatcher, H., Stantchev, R.I., Wang, J., Li, K., Hernandez Serrano, A., Taylor, Z.D., Castro-Camus, E. and Pickwell-MacPherson, E., 2022. Terahertz (THz) biophotonics technology: Instrumentation, techniques, and biomedical applications. *Chemical Physics Reviews*, 3(1). aip.org
- Balasundaram, G., Krafft, C., Zhang, R., Dev, K., Bi, R., Moothanchery, M., Popp, J. and Olivo, M., 2021. Biophotonic technologies for assessment of breast tumor surgical margins—A review. *Journal of biophotonics*, 14(1), p.e202000280. researchgate.net
- Fanous, M.J., Casteleiro Costa, P., Işıl, Ç., Huang, L. and Ozcan, A., 2024. Neural network-based processing and reconstruction of compromised biophotonic image data. *Light: Science & Applications*, 13(1), p.231. nature.com
- Verma, A. & Gupta, S. K., 2023. History and Techniques of Bioimaging. *Magnetic Quantum Dots for Bioimaging*. [HTML]
- Jourlin M., Carré M., Breugnot J. , Bouabdellah M., 2012, *Logarithmic Image Processing: Additive Contrast, Multiplicative Contrast, and Associated Metrics*, Editor: Peter W. Hawkes, *Advances in Imaging and Electron Physics*, Elsevier, Volume 171, 2012, Pages 357-406, ISSN 1076-5670, ISBN 9780123942975.
- Shannon C.E. 1948, A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), pp. 379-423.
- Golan, A. & Foley D.K., 2023. Understanding the Constraints in Maximum Entropy Methods for modeling and Inference, *IEEE Trans Pattern Anal Mach Intell.* 2023 Mar;45(3):3994-3998. doi: 10.1109/TPAMI.2022.3185394. Epub 2023.
- Golan, A., 2018. *Foundation of Info-Metrics: Modeling, Inference, and Imperfect Information*. New York, NY: University Press, pp. 1-465.
- Golan, A., 2008. *Information and Entropy Econometrics. A review and synthesis*, *Foundations and Trends in Econometrics*, 2, pp. 1-145.
- Golan, A., Judge, G. and Miller, D., 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, New York, John Wiley & Sons.
- Bernardini Papalia R. and Fernandez Vazquez E., 2021, *Forecasting Socioeconomic distributions on Small-Area Sampling Domains for count data*, in *Advances in Info-Metrics: Information and Information Processing across Disciplines*. Eds: M. Chen, J.M. Dunn, A. Golan, A. Ullah, Oxford University Press, pp. 240, 263.
- Bernardini Papalia R. and Fernandez Vazquez E., 2018, *Information theoretic methods in small domain estimation*, *Econometrics Reviews*, 37, 4, 347-359.
- Bernardini Papalia R., 2024, *Modeling the Learning from Repeated Samples: a Generalized Cross Entropy Approach*, working paper. Bernardini Papalia R., Gullà D., Nastati E., 2023, *Quantum agriculture and experimental detection of wheat flour quality using thermal image technology*, *Heliyon*, Volume 9, Issue 9, e19899

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.