Extracting the most relevant information from biophotonic data: a Constrained Maximum Entropy methodology

ROSA BERNARDINI PAPALIA
Department of Statistical Sciences,
University of Bologna,
ITALY

Abstract: This paper presents a novel approach, the Constrained Maximum Entropy (CME) methodology, for extracting knowledge from biophotonic data. More specifically, we discuss the main issues related to this new type of data and demonstrate the potential of the CME methodology to incorporate both *a priori* knowledge and data constraints to efficiently analyze biophotonic data. The key advantage lies in its ability to determine the most "unbiased" biophoton distribution - one with maximum entropy among distributions that satisfy given constraints while remaining uncommitted to unavailable information. Furthermore, we advance the discussion by proposing that the CME formulation, enriched with quantitative and qualitative constraints derived from precise biophoton emissions, serves as a powerful new tool for monitoring changes in biological systems. It holds the potential to identify unstable states and assess the impact of novel treatments on these systems. An empirical application for the plant study based on imaging sensors and AI mathematical algorithms is also provided.

Key-words: Biophotons emission, Quantum Data, Constrained Maximum Entropy.

Received: March 6, 2025. Revised: June 3, 2025. Accepted: June 21, 2025. Published: July 15, 2025.

1. Introduction

Although advanced data acquisition technology is widely available, extracting the most relevant information from biophotonic data for efficient application across diverse fields remains a sophisticated problem to address. The sheer volume of data, coupled with its complexity and the time required for interpretation, presents significant obstacles. There is a natural tendency to limit information to what is perceptible to the human eye. However, biophotons offer a wealth of valuable insights that can only be accessed through rigorous quantitative analysis. The scattering of light within biological systems contributes to what is called polarized light. By measuring the polarized properties of light, constructing models, and simulating its propagation, we can unveil intricacies far too complex to be interpreted through qualitative means alone. The development of inverse algorithms and models presents a promising avenue for extracting meaningful insights. An inverse problem involves determining the cause based on an observed effect. In the context of this research, the goal is to ascertain the optical properties of a biological sample and their correlation to specific conditions or diseases. This requires comparing measured data against a model, accompanied by efforts to minimize discrepancies between the model and actual observations. Given the complexity of the data and the underlying optical properties, inverse algorithms must be carefully designed to address these intricacies. Biophotonics often encounters challenges such as ill-defined matrices and non-linear systems. The first step in tackling an inverse problem is always simplifying it to a manageable and practical form. When executed effectively, the extraction of relevant data can serve as a powerful tool for general decision-making and more intricate evaluations relying on non-invasive, light-based techniques.

Harnessing the power of visualization, measurement, and quantification at micro and nano scales, biophotonics has made remarkable contributions to discoveries and advancements in life sciences. Inevitably, the complexity and volume of data generated by various biophotonic imaging and sensing techniques have grown exponentially, underscoring the critical need for efficient data analysis to convert raw information into actionable knowledge. It is essential to ensure that the data is not interpreted solely within the context of the acquisition method, hypothesis, or expected outcome, as this risks introducing biases. Instead, the focus should be on deriving the best possible

answer to the initial inquiries, laying a robust foundation for knowledge expansion with minimal assumptions. Through precise and reliable analysis, biophotonic data can evolve into trusted information - information capable of driving new discoveries and shaping the future of research. The significance of data analysis in biophotonics is further highlighted by its direct relatability to human health and the well-being of biological systems (animal, plants…). Ultimately, the trusted insights gleaned from this data will profoundly inform decision-makers and policy creators.

The objectives of this work are: (i) to provide a method for the identification of the most crucial problems in biophotonic 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 biophotonic researcher in getting the most out of their experiments and data analysis. An empirical application for the plant study based on imaging sensors and AI mathematical algorithms is also provided.

2. Methods and experimental framework

2.1 Background of biophotonic data

Real biophotonic data sets are derived from the analysis of light propagation in biological systems. The data sets are typically extensive, ranging from 128x128 to 512x512 or larger, and are presently analyzed point by point, yielding information solely on scattering. This underscores the primary and most critical challenge in biophotonic data analysis, that of developing an understanding of the various factors contributing to light transport within biological systems. Light transport in tissues is a complex process involving absorption, scattering, fluorescence, and luminescence. Developing a comprehensive model of light transport in tissues and how it can be best characterized by specific types of measurements is an unresolved challenge. Furthermore, the absence of a unified framework to interpret biophotonic data in a biological context adds another layer of complexity to data analysis in this field, [1], [2], [3], [4], [5], [6], [7], [8], [9].

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 unique challenge specific to biophotonic data analysis is its interdisciplinary nature; physicists or engineers conducting measurements may lack a strong biological background, while biologists or medical researchers may struggle to characterize the data from a physical perspective. This type of data is becoming more prevalent with the growing popularity of purely optical clinical measurements and is typical of research data that often blurs traditional disciplinary boundaries. Interdisciplinary efforts face inherent challenges, such as communication barriers between researchers from diverse fields and the difficulty of interpreting data outside one's specific area of expertise. This can cause difficulties when biophotonic data is analyzed or interpreted by someone who does not have a strong understanding of its biological significance.

Biophotonic data refers to information obtained and stored in digital form from biological samples through optical methods. This encompasses a variety of data types, including microscopic and spectroscopic images, spectra, and information derived from light-based therapeutic interventions. Key characteristics of biophotonic data include its ease of exchange and adaptability for analysis at locations distinct from the point of acquisition thanks to its digital nature.

Typically non-invasive, biophotonic digital optical methods can be used to monitor the effects of treatments in longitudinal studies while also enabling the immediate application of new approaches to biological systems. Moreover, as biophotonic data is obtained using optical means, it can often be readily interpreted even by those outside the specialized field of research.

Effective communication between researchers across diverse disciplines is essential, as it enables the exchange of knowledge and methodologies, fostering interdisciplinary collaboration. Most importantly, there is immense potential for advancing the treatment of unstable conditions through the application of methods informed by the analysis of biophotonic data.

2.2. Characteristics of biophotonic 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 often involves measuring signaling events within 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.

Biophotonic 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 biophotonic 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 integrates prior knowledge about the system, such as established physical principles, and generates predictive outcomes based on that framework.

Spectral parameters relate to the fact that photons can have different energy sources. 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 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 provide information about the depth of light penetration into the tissue. A notable example of biophotonic data with spectral resolution includes the reflectance or fluorescence spectrum obtained from specific tissues

Spatial and temporal parameters are among the most intuitive to grasp, as they correspond to the location and timing of the data acquisition

process. Temporal information is particularly significant in cases where photons are measured over a span of time, such as in fluorescence lifetime measurements.

2.3. Pre-processing techniques for biophotonic data

Pre-processing routines are essential prior to data modeling, as they help to reduce noise and ensure reliable input for subsequent analysis. Noise refers to any data that is uninterpretable, which in biophotonic imaging is often attributed to low-light conditions and auto-fluorescence from the sample. Noise usually arises from biophotonic data due to variations in the data capture process. Imaging techniques can be implemented to significantly reduce specific types of noise, although this may inadvertently introduce other noise forms. Signal processing aims to isolate the signal of interest by distinguishing it from noise. Images can degrade due to additive or multiplicative noise, and the goal is to remove noise from the image while preserving 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, which can complicate its interpretation.

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 refine it into a cleaner form. These methods often rely on minimal assumptions about the data and employ established statistical and mathematical theories to estimate the true signal. 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. The success of noise removal can be evaluated through statistical analysis and by comparing the processed, noise-free data against the original dataset to perform an error assessment.

A low-pass filter is commonly employed in noise removal, allowing image components with frequencies below a specified cut-off frequency to pass through while attenuating those with higher frequencies, [10]. 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, along with certain types of random noise, can be effectively reduced through wavelet transformation, which decomposes a signal into a set of basic functions for analysis. Each basis function has a different level of localization in time and frequency. Wavelet thresholding offers a distinct advantage over Fourier analysis since data and functions can be analyzed across multiple scales. By applying a threshold to the wavelet coefficients, the noisy coefficients can be removed, and the signal can be reconstructed with reduced noise. Maximum Entropy thresholding appears to be a highly effective method for segmenting fluorescence lifetime images. Its key advantage lies in its foundation on a consistent statistical framework, ensuring reliability and precision in segmentation. 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 straightforward, this method is generally too restrictive, as it assumes that the noise present in the signal and the baseline are equal. In practice, this is often not the case. The true objective is to compare the signal to a noise-free baseline, which does not necessarily need to share the same shape as the signal baseline. Nevertheless, the method offers the advantage of enabling comparisons between images acquired under varying conditions. Additionally, it is reversible, allowing the

transformed image to be compared directly to the original data.

Traditional fitting methods, including robust fitting, component model separation, wavelet analysis, and singular value decomposition, perform adequately in areas where model simplification may sufficiently describe observed behaviors. However, these methods fail to incorporate learned information, which can be represented using 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. Semi-classic and intrinsic Bayes classifiers excel in unsupervised and supervised classification tasks, provided Gaussian assumptions are met, offering a more robust approach to data analysis under such conditions.

2.4. Constrained Maximum Entropy methods in biophotonics data analysis

The advancement of experimental biophotonics methods is accompanied by a pressing demand for efficient tools and techniques to analyze the resulting data effectively 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. Integrating innovative tools and methodologies from other scientific disciplines into biological research has emerged as a prominent trend in modern 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.

Maximum Entropy based methods grew out of the development of the principle of maximum caliber in the 1950s. 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, [11].

This result complements other fundamental results of statistical mechanics but also generalizes them substantially. Although efficient measurements are available for a limited subset of properties and rare system states, the untapped potential within the remaining microstates holds valuable information. This potential gives rise to alternative, stable, and often qualitatively robust probability distributions that may significantly deviate from the equilibrium distribution. Utilizing that information serves as an effective approach to developing novel, testable hypotheses, enabling the evaluation of statistical models and their confidence levels while minimizing undue bias. Moreover, mutual information stands out as a particularly compelling association measure for high-throughput studies of intracellular networks, thanks to its information-theoretic properties, broader scope, and capacity to uncover nonlinear relationships even in noisy environments. It has been applied in both experimental and computational studies across a variety of systems, ranging from simple preservation to more intricate scenarios. In such cases, traditional methods may be prone to errors due to the challenges posed by multiple hypotheses, especially when experimental control samples are lacking, the range of comparisons is limited, and the dataset's high dimensionality leads to significantly skewed feature distributions. Key challenges for the effective application of the mutual information measure include achieving rapid and accurate association estimation, maintaining acceptable control over false associations, and ensuring sufficient noise robustness in association detection. Numerous communication and information-theoretic function approximations, as well as limit theorems, have demonstrated that the entropy functional (and the information functional in conditional cases) exhibits Gaussian self-averaging properties in the large sample size limit. With the increase of sample size, the dependence of the associated estimation on specific estimation algorithms tends to diminish.

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, [12], [13], [14], [15], [16], [17]. The theory of maximum entropy learning assumes a set of mappings from observed events to constraint features.

The challenge of estimating an unknown probability function p from a set of observed data X1, …, Xn is a cornerstone of statistical analysis. Here, we focus on a relatively intricate feature space and adopt a highly non-parametric modeling approach.

Although prior knowledge regarding X, which we group as features f1, ..., fj, may be available, the only information allowed in the CME model formulation is a list of constraints. When accurate, this information can significantly enhance the learning of a better posterior. However, even in cases where the observation space X is relatively simple, leveraging such potentially useful information without risking overfitting the hypothesis class remains a more complex and nuanced challenge. 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 other methods involving more basic assumptions (as normality and/or lack of correlation of the data generation process).

2.5. Constrained Maximum Entropy (CME) formulation for detecting biophotonic distributions

One of the major challenges in the field of biophotonics is systematic data analysis and the extraction of knowledge from the generated data. To address this issue, general maximum entropy methods—commonly implemented through the negative Laplace transform or various inversion techniques—are proposed for analyzing data derived from light transmission across a defined set of spatial sampling points.

Furthermore, general maximum entropy methods, often framed as weak-constraint minimization problems, can be applied to space-time hyperspectral imaging of biological systems. These methods, combined with a new class of nodes and compressive sampling, enable the extraction of substantially more information about labels from the same number of measured diffraction patterns compared to conventional techniques, outperforming what compression sampling alone can achieve.

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 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 accommodated.

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-suited to addressing the intrinsic fuzziness and associated estimation variability of biophotonic data. They also appear to account for the quantum uncertainty inherent in the collection of biophotonic image data, enabling a

comprehensive statistical description of both the data modeling results and their estimation uncertainty.

The availability and advantages of this additional methodological support complement the deep dictionary learning approach used in advanced experiments, helping to bridge the gap between the data and the sophisticated 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 meets the given constraints, maximizes Shannon information entropy or minimizes Kullback-Leibler divergence, and remains least biased toward information that is not yet available. Additionally, building on the works of [13] and [16], this framework can be extended to incorporate noise in the constraints, reflecting the uncertainty surrounding the process under study and the nature of the variable's photons count.

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

$$c_i = p_i + \varepsilon_i \tag{1}$$

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:

$$c_i = \sum_{m=1}^{M} p_{im} z_m \tag{2}$$

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

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

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 $\boldsymbol{v}' = \{v_1, \dots, v_L\}$, 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:

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

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):

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

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

The solution to the estimation problems is given by minimizing the KL divergence between the $\boldsymbol{p}'$s and

the $\boldsymbol{q}$'s. Specifically, the constrained minimization problem can be written as:

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

Subject to:

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

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

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

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 ($\boldsymbol{q^{\alpha}},\boldsymbol{q^{\beta}},\boldsymbol{W^0}$) are specified as uniform, then the GCE solution reduces to the GME one.

To recover the probability vectors $\boldsymbol{p_i}$, the Lagrangian Function following the matrix form can be expressed as follows:

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

with the first order conditions:

$$\frac{\partial \mathrm{L}}{\partial p_{im}}=\sum_{i=1}^{T}(\ln(\tfrac{p_{im}}{q_{im}})+1)-\lambda\sum_{m=1}^{M} z_m-\sum_{i=1}^{T}\sum_{m=1}^{M}\vartheta_i=0$$

$$i=1,\dots,T; m=1,\dots,M$$

$$\frac{\partial \mathrm{L}}{w_{il}}\sum_{i=1}^{T}(\ln(\tfrac{w_{il}}{w_{il}^0})+1)+\lambda\sum_{l=1}^{L} v_l-\sum_{i=1}^{T}\sum_{l=1}^{L}\mu_i=0$$

$$i=1,\dots,T; l=1,\dots,L$$

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

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

$$\frac{\partial L}{\partial \mu_i}=1-\sum_{l=1}^{L} pw_{il}=0; l=1,\dots,L \quad (9)$$

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

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

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.

3. An application to Plant Studies through Computational Photons View

Biophotons are emitted by living cells when their molecules transition from an electronically excited state to a stable ground state, resulting in the spontaneous release of ultra-weak luminescence linked to biochemical processes. So far, the basic molecular mechanisms underlying biophoton emission are unknown. To analyze the behavior of photons in the environment (flow, directionality, reflectance, absorbance, and scattering), a basic method setup is required. The scattering biophoton reading effectively captures the finest variations in light intensity, vibrational and spectral components, and can detect changes in the quantum characteristics of biological materials.

The first phase of the applied study was to record images of the flour sample and read photometric data from luminescent images by means of the HIS software. The images were generated by webcam and the bio-photonic emissions of the two wheat flour samples measured using imaging software Vibraimage HIS are analyzed to estimate the photonic frequency distribution using the CME proposed method. After capturing the images, mathematical analyzes of the pixel's variation were performed, to evaluate signs in plants, based on the images, algorithms found in VibaHIS. The generated images were collected in AVI format (Audio and Video Interlayer file) with 1 frame per second (fps). Several numerical parameters of this sample were acquired, and the average was computed on the image stack. Following the image capture, mathematical analyses of pixel variations were conducted to assess signals in plants using algorithms implemented in VibaHIS. Multiple numerical parameters were extracted from these samples, and their averages were calculated across the image stack.

The preliminary photon analysis was conducted using samples of the Ludwig variety organic flours (a quantum treated flour following Nastati protocol, [18]. The analysis was based on the results produced in terms of the variability of biophotonic frequencies (emissions/signals) and consistency indices using the Vibraimage HSI Software.

The second phase of this study was finalized to identify the shape of the biophotonic frequency distribution that accounts for the quantum uncertainty inherent in the collection of biophotonic image data, utilizing the proposed CME method (Fig.1). The shape of the distribution of the biophotons emission detected in our experiment seems to be asymmetric and it is impossible to obtain a good of fit of it using either Poisson or other methods based on the Normality assumption.

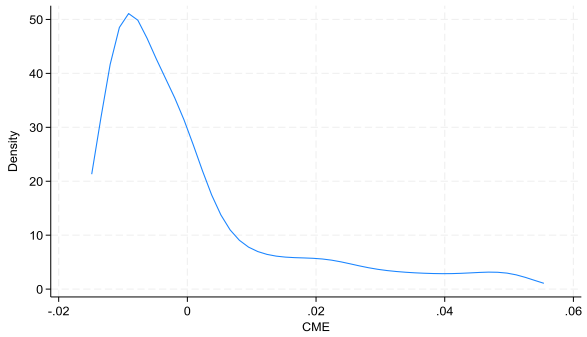

Fig. 1: CME based density distribution of biophons, flour Ludwig variety with Nastati treatment

The final analysis was based on the comparison between asymmetrical Kernal the Log Normal Kernel based count function, (Fig.2) [19], and the CME based distribution function (Fig.1). Matlab and R packages can be applied to a wide range of optimization tasks, including continuous, discrete, mixed and constrained optimization problems with reference to a framework for non-convex optimization using the Cross-Entropy [20], [21].

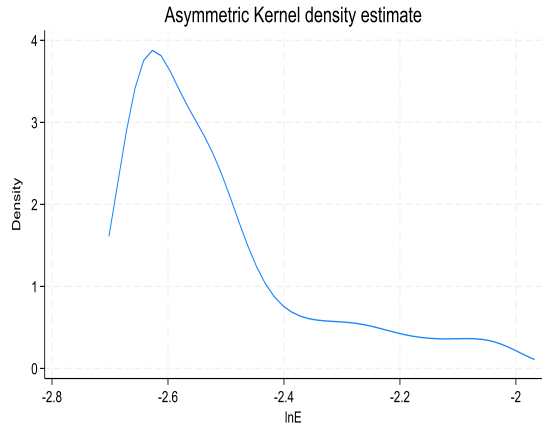

Fig. 2: Log Normal Kernel density distribution of biophons, flour Ludwig variety with Nastati treatment

The asymmetric log normal Kernel estimator is used to account for asymmetry of the probability distribution function of biophotons counts. The results showed how there is clear evidence that biological systems cannot be described by the ordinary prescriptions of equilibrium statistical mechanics. The need to consider signals of very low intensity coming from nonstationary processes emerge.

Our results provide an empirical basis for further extending this study to predict other important determinants of wheat quality, associated with the statistical properties of the distribution, using the imaging technology setup. The computer vision technology integrated into VibraHSI software offers versatile support for laboratory and agronomic fieldwork, utilizing dynamic images captured by various types of video equipment.

The use of images, sensors and mathematical algorithms can help in the generation of technical attributes and facilitate the plant health diagnosis. Combined with this, computer vision provides a non-destructive and non-invasive strategy for collecting samples and analyzing plant treatments.

Further experiments are required to replicate the results of this study under different environmental conditions. In addition, suitable comparative chemical analyses should improve the knowledge of the properties of the biophotons count distribution that are very important to identify useful biological information.

4. Conclusion and Perspectives

In this work, we have proposed an IT based method to extract relevant information from biophotonic data. Photonic data represents a map of light intensity and scattering from a sample. Such data, generated through diverse experimental setups, is exceptionally rich in information and holds immense potential to deepen our understanding of nature, the equilibrium of biological systems, and the changes that affect them. The objective is to develop automated techniques to interpret this novel type of data effectively.

The field offers vast opportunities for future research, given the variety of light scattering experiments and the range of potential biological compositions and structural changes in systems. 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 crucial for leveraging the full potential of such data to tackle intricate questions across diverse experimental contexts. There is potential to develop these methods into a practical tool for applied researchers; this requires the development of inverse algorithms to estimate pattern difference maps from real data, along with validation against experimental studies on biological systems.

Scattering patterns from well-characterized changes can be effectively quantified relative to normal states through the use of pattern difference maps. These maps display variations in scattering intensity at each spatial point, comparing unbalanced or altered states to normal, stable states. It is possible to predict such pattern difference maps computationally with a good degree of accuracy using an approximation of light transport. This is important because it provides a forward model for interpreting pattern difference maps. Pattern difference maps can ultimately be transformed into images of change that vividly depict alterations in biological structures. Such images have the potential to offer groundbreaking insights into structural changes within biological systems and their connections to 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. This innovative approach extracts the most significant changes from biophotonic data and facilitates the aggregation of this information through traditional statistical analyses. The method is widely applicable to various challenges in which changes in complex multivariate data need to be detected and interpreted, representing a general advance in the analysis of biophotonic data. Notably, it is particularly effective for analyzing spectral data obtained from optical measurements of biological systems.

The results found in the plant study based on imaging sensors and AI mathematical algorithms demonstrate the efficiency of image analysis technology as a great tool to support the agronomic fieldwork. The proposed method seems to provide useful information on the parameters of the biophotons count distribution without imposing any strong assumption on the generating data process. The advantage is that this method requires less restrictive assumptions on data and can determine the statistical properties of the light/photons considering that there is a direct correspondence between the functional form of the estimated photon distribution and the statistical properties we are looking for. These properties are the final objective of the statistical analysis because they provide useful information on the biological state of the plant.

In summary, our information-based approach offers a versatile tool for designing and evaluating diverse imaging systems, as demonstrated across various applications. The method's generic probabilistic foundation suggests potential applications beyond imaging, extending to electronic, biological, geological, and chemical sensors.

Future research in data mining from optical measurement systems offers numerous promising directions. One particularly impactful approach involves data fusion from multiple sources. For instance, biological system samples measured optically are often analyzed using additional instruments that assess chemical properties. Integrating data from these devices with optical measurement data could significantly enhance the insights gained from the original samples. To facilitate data fusion, it is essential to establish measurement standards in biophotonics, enabling meaningful comparison and integration of data from different instruments and laboratories. Addressing this challenge requires collaboration among interdisciplinary teams. Furthermore, the fusion of data and estimation through entropy maximization holds the potential to incorporate diverse types of information, even at an aggregate level.

References

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

[2] Pan, T., Lu, D., Xin, H., & Li, B., 2021. Biophotonic probes for bio-detection and imaging. *Light: Science & Applications*. nature.com.

[3] 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.

[4] 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]

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] Verma, A. & Gupta, S. K., 2023. History and Techniques of Bioimaging. *Magnetic Quantum Dots for Bioimaging*. [HTML]

[10] 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.

[11] Shannon C.E. 1948, A Matematical Theory of Communication. *Bell System Technical Journal*, 27(3), pp. 379-423.

[12] 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.

[13] Golan, A., 2018. *Foundation of Info-Metrics: Modeling, Inference, and Imperfect Information*. New York, NY: University Press, pp. 1-465.

[14] Golan, A., 2008. Information and Entropy Econometrics. *A review and synthesis, Foundations and Trends in Econometrics*, 2, pp. 1-145.

[15] Golan, A., Judge, G. and Miller, D., 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, New York, John Wiley & Sons.

[16] 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.

[17] Bernardini Papalia R. and Fernandez Vazquez E., 2018, Information theoretic methods in small domain estimation, *Econometrics Reviews*, 37, 4, 347-359.

[18] 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.

[19] Khan J.A., Akbar A., Golam Kibria B.M. 2023, DELTD: An R Package for Kernel Density Estimation using Lifetime Distributions, *WSEAS TRANSACRIONS on Mathematics*.

[20] A. Cunha Jr, M. V. Issa, J. C. Basilio, and J. G. Telles Ribeiro, *CEopt: A MATLAB Package for Non-convex Optimization with the Cross-Entropy Method*, PrePrint, arXiv:2409.00013, 2024, DOI https://doi.org/10.48550/arXiv.2409.00013.

[21] Benham T., Duan Q., Kroese D.P., Liquet B. (2017) CEoptim: Cross-Entropy R package for optimization. *Journal of Statistical Software*, 76(8), 1-29.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The author contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The author has no conflict of interest to declare that is relevant to the content of this article.