2017

Enhanced optical alignment of a digital micro mirror device through Bayesian adaptive exploration

Kevin H. Knuth
University at Albany, State University of New York

Kevin B. Wynne
University at Albany, State University of New York

Jonathan Petruccelli
University at Albany, State University of New York

Enhanced optical alignment of a digital micro mirror device through Bayesian adaptive exploration

Cite as: AIP Advances **7**, 125207 (2017); https://doi.org/10.1063/1.5002539
Submitted: 30 August 2017 • Accepted: 29 November 2017 • Published Online: 08 December 2017

Kevin B. Wynne, Kevin H. Knuth and Jonathan Petruccelli

Enhanced optical alignment of a digital micro mirror device through Bayesian adaptive exploration

Kevin B. Wynne,[1,2] Kevin H. Knuth,[1] and Jonathan Petruccelli[1]
[1]*Department of Physics, State University of New York Albany, Albany, NY 12222, USA*
[2]*Department of Physics, Hudson Valley Community College, Troy, NY 12180, USA*

(Received 30 August 2017; accepted 29 November 2017; published online 8 December 2017)

As the use of Digital Micro Mirror Devices (DMDs) becomes more prevalent in optics research, the ability to precisely locate the Fourier "footprint" of an image beam at the Fourier plane becomes a pressing need. In this approach, Bayesian adaptive exploration techniques were employed to characterize the size and position of the beam on a DMD located at the Fourier plane. It couples a Bayesian inference engine with an inquiry engine to implement the search. The inquiry engine explores the DMD by engaging mirrors and recording light intensity values based on the maximization of the expected information gain. Using the data collected from this exploration, the Bayesian inference engine updates the posterior probability describing the beam's characteristics. The process is iterated until the beam is located to within the desired precision. This methodology not only locates the center and radius of the beam with remarkable precision but accomplishes the task in far less time than a brute force search. The employed approach has applications to system alignment for both Fourier processing and coded aperture design. https://doi.org/10.1063/1.5002539

I. INTRODUCTION

Spatial light modulators (SLMs) are widely used in optical research to modulate light in a precise and controllable manner. They have been used for beam shaping,[1–7] computational imaging,[8,9] the design of compact high-dynamic range cameras[10,11] and numerous other areas of research where manipulation of the spatial distribution of light is beneficial such as polarization,[12] pulse shaping.[13]

The application explored here utilizes an SLM as a Fourier plane filter. A Fourier plane is a plane in an optical system where the Fourier transform of an input field exists. Points in the Fourier plane are the spatial representation of the light's direction, and by placing a mask in the Fourier plane these directional components can be manipulated: their amplitudes can be modulated or relative phase delay can be introduced between them. For example, consider a sample illuminated with a collimated beam (one made of parallel rays). These parallel rays will converge to a bright point at the center of the Fourier plane, rays scattered away from parallel will arrive at the Fourier plane some distance from this central point. By introducing a mask into the system that obstructs the central bright spot, a dark field image is created in which features of the object that have scattered light will be imaged brightly against a dark background. Zernike phase contrast methods also employ Fourier plane masks to enhance scattered light through manipulation of both its amplitude and phase.[14]

When using multiple Fourier filters, the ability to dynamically vary the Fourier mask without the need for mechanical switching or scanning of filters is advantageous since it obviates the need for mechanical registration and can be done rapidly. Broadly speaking, one type of technology commonly used for this purpose is a liquid-crystal-based SLM that can rotate the polarization of incident light at the pixel level, either in reflection or transmission. With the appropriate polarizing filters, such SLMs can allow modification of phase and/or amplitude of the light. Although these methods are dynamic and useful, the extinction ratio (contrast) is not 100 percent, frame rates in high resolution commercial SLMs are limited (often to 60 Hz) and it can be difficult to precisely control only the

amplitude without impacting the phase. Moreover, an SLM will typically perform as desired only for a single wavelength of light.

Another SLM technology is the digital micro-mirror device (DMD) which operates by selectively tilting tiny mirrors between two orientations to control the reflection of the incident beam. If collecting optics are positioned to receive light from the DMD when it is in one of the two states a true pixel-by-pixel binary control is created. DMDs allow rapid (kHz and beyond) refresh rates and work over a broad spectrum. However, they allow only binary amplitude modulation of the field.

One challenge when using SLMs as dynamic Fourier filters is precisely locating the position of the projected pattern precisely with respect to the center of the beam's Fourier transform. Although this could be accomplished by other methods, Bayesian adaptive exploration offers an approach that is well suited to this type of task. While the focus here is limited to DMDs, the methods explored can be readily adapted to other SLMs.

The physical setup, which will be discussed in detail in the experimental design section, allows DMD pixels in the Fourier plane to be directed towards or away from a camera or power meter, which then acquires an intensity measurement. The information gained by this measurement is then used to estimate the parameters describing the beam's size and location in the Fourier plane using a Bayesian inference engine. A Bayesian inquiry engine is then used to choose the next location to measure based on the maximization of the expected information gain.

Details of this method will be discussed at length in the Computational Methodology section. The algorithm works through an iterative process switching DMD pixels on and off while taking power measurements in order to determine the size and location of the beam with high precision using the least amount of data possible.

A. Fourier optics

Neglecting polarization, a monochromatic beam of light can be represented as a complex field $U(r)$ where r denotes position. This can be expressed as the superposition of individual plane waves each possessing a real amplitude a_n, phase delay δ_n, a dimensionless unit vector pointing in the direction of propagation of the wave $\hat{\mathbf{u}}_n$, and a wave number $k = 2\pi/\lambda$, where λ is the wavelength of the beam of light. Although a field would typically be represented by an integral over all possible plane wave directions, a discrete set of plane waves will be assumed, and a summation used for notational simplicity

$$U(\mathbf{r}) = \sum_{n=1}^{N} a_n e^{\mathrm{i}[k(\hat{\mathbf{u}}_n \cdot \mathbf{r}) + \delta_n]}. \tag{1}$$

As the beam of light passes through, or reflects off of an object the individual plane waves are perturbed. This perturbed field contains amplitude and phase components that would provide detailed information about the object itself, with the amplitude describing the absorption of light and the phase describing the thickness/depth.

However, light in the visible spectrum oscillates on the order of $4 - 8 \times 10^{14}$ Hz which is too rapid for the field to be directly measured with an imaging camera. As a result, only the time-averaged amplitude or intensity I of the electric field is detected at any point. The intensity can be modeled as

$$I = |U|^2 \tag{2}$$

and is defined as the power per unit area striking a surface oriented perpendicular to the beam's direction of propagation. Note that typically only relative variations in intensity matter and imaging cameras produce a measurement proportional to power per unit area (the exact proportionality constant being unimportant). As a result of multiplying the field by its complex conjugate, the information about the phase and direction of the complex field is lost. However, the Fourier filtering of a field can allow its phase to be rendered as measurable variations in intensity.

A lens has the remarkable quality of performing a Fourier transform of the optical beam passing through it. A simple system for Fourier filtering can be built by placing two identical lenses of focal length f at a separation of $2f$ from each other, this results in a configuration known as a $4f$ system illustrated in Fig. 1. Consider an input object at Σ_o illuminated by a normally incident plane wave, such that the perturbed field from the object consists of a collection of plane waves with different

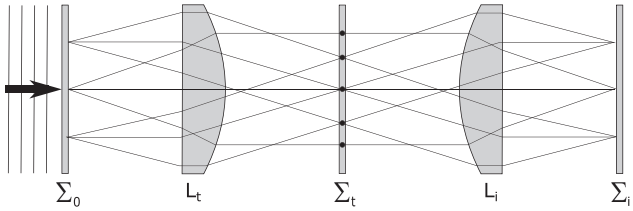

FIG. 1. Graphical representation of a 4f system of lenses.

amplitudes and phases. The first lens L_t produces a Fourier transform of this field at the midpoint between the lenses, Σ_t. The second lens L_i undoes this Fourier transform, producing an inverted copy of the field at Σ_i. The importance of this type of optical system is that it creates a Fourier plane between the lenses that can be manipulated by physical means in order to create a Fourier domain filter.

The Fourier transformation itself creates a map of the spatial frequencies of individual plane waves. Consider a collection of individual plane waves represented by the field U leaving the object Σ_o. Each plane wave passes through the lens L_t in different directions $\hat{\mathbf{u}}_n$, these directions translate to different spatial frequencies directing each plane wave to a specific point at the Fourier plane Σ_t. Plane waves traveling parallel to the optical axis converge to a point at the center of the Fourier plane, and those with larger angles converge at points further from its center. The pattern that manifests itself at the Fourier plane Σ_t is transformed back to the original complex field U after passing through the second lens L_i. The grouping of spatial frequencies at the Fourier plane allows selective inclusion or exclusion of individual plane waves contained within the complex field if a mask is placed in that plane.

At this point, it might be instructive to consider a crude technique utilizing Fourier filtering that would recover the phase information lost to time averaging. If only two of the many plane waves that make up the complex field could be isolated, for example by placing two pinholes in the Σ_t plane, the phase difference between the two waves could be recovered. The intensity distribution in the image plane Σ_i is then given by

$$I = a_1^2 + a_2^2 + 2a_1a_2 \cos\{k\,[(\hat{\mathbf{u}}_2 - \hat{\mathbf{u}}_1)\cdot\mathbf{r}] + (\delta_2 - \delta_1)\} \tag{3}$$

where the subscripts 1 and 2 indicate the two plane waves allowed through the Fourier domain mask. Equation (3) describes the sum of the individual intensities a_1^2 and a_2^2 of the two waves and the sinusoidal pattern resulting from the interference of the two waves.

If the intensity were measured with only one pinhole opened at a time, a_1^2 and a_2^2 could be directly measured. A measurement with both pinholes opened as in Eq. (3) would then allow the phase difference to be calculated from the resulting interference pattern. This can be readily achieved with a DMD placed in the Fourier plane by turning on only two pixels on the DMD, individually and then both together. This method could be extended to include all pairwise combinations of available plane waves to determine the phase of the complex field, however this brute force approach is not practical as it is highly inefficient, using all the detector pixels in three separate measurements to obtain three scalar parameters and would require 2N-1 measurements to retrieve the phase differences for N plane wave components of the field. More sophisticated algorithms exist that could use the DMD as a Fourier filter to accomplish the same results more efficiently.[15,16] Because knowing which plane waves correspond to which DMD pixels is critical for these methods, it is important to precisely locate the DMD within the Fourier plane. It is for this reason that the focus of the process developed here is to assure this precise location.

Note also that Eq. (3) describes an intensity across the camera that is sinusoidal interference pattern whose wavelength is $\lambda/|\hat{\mathbf{u}}_2 - \hat{\mathbf{u}}_1|$ such that widely separated plane wave directions are required to resolve sharp features at the camera. As a consequence, Fourier plane filters can also modify the system's resolution. By passing only light near the center of Fourier space, only large intensity features can be produced at the camera, reducing the resolution of the image. Furthermore, due to the directional nature of the interference pattern in Eq. (3), blocking the outer regions of the Fourier

plane in the horizontal direction reduces horizontal resolution and in the vertical direction reduces vertical resolution.[17]

B. Computational methodology: Inquiry and inference engines

The DMD consists of an array of ten micron square micro mirrors that can be individually addressed and operated through Matlab. The mirrors exist in one of two states, they can tilt by twelve degrees in a positive or negative direction relative to an axis parallel to the surface of the DMD. Simple experiments can be performed that consist of engaging a micro mirror or a set of neighboring micro mirrors and measuring the reflected intensity. Each experiment has the potential to provide some information about the position and radius of the beam on the DMD.

The beam itself is modeled as a circle projected onto the planar surface of the DMD. This circle is summarized by three model parameters: the x-position of the beam center x_0, the y-position of the beam center y_0, and the beam radius r_0. The goal is to learn the values of these model parameters that best describe the beam on the DMD. This is accomplished by relying on the coupling of two software engines: an inquiry engine, which decides which experiments to perform by flipping micro mirrors, and an inference engine that learns all it can about the beam parameters from the data collected from all of the experiments.

The inference engine relies on applying Bayesian probability theory along with all of the data collected in a set of experiments to infer the three beam parameters. This is accomplished by employing a Monte Carlo sampling technique known as nested sampling,[18] which explores the posterior probability density of the three beam model parameters.

The inquiry engine utilizes samples drawn from the posterior probability density of the beam model parameters to determine a set of optimal micro mirror locations on the DMD so that flipping those micro mirrors would provide the greatest expected information gain to aid in the search for the beam's circle. In essence, the inquiry engine designs an optimal experiment to perform based on the information provided by the proceeding experiments.

Although the computational details underlying the inference and the inquiry engines are described in the following sections, an overall description of the algorithm structure is represented in Fig. 2. The algorithm begins by flipping on all of the mirrors of the DMD one quadrant at a time in an attempt to narrow down the search to one quadrant. It then records a uniformly distributed set of intensity measurements from within that quadrant to seed the inference engine for the first iterative loop. These initial data points are not required for the algorithm to converge,[27,28] but they

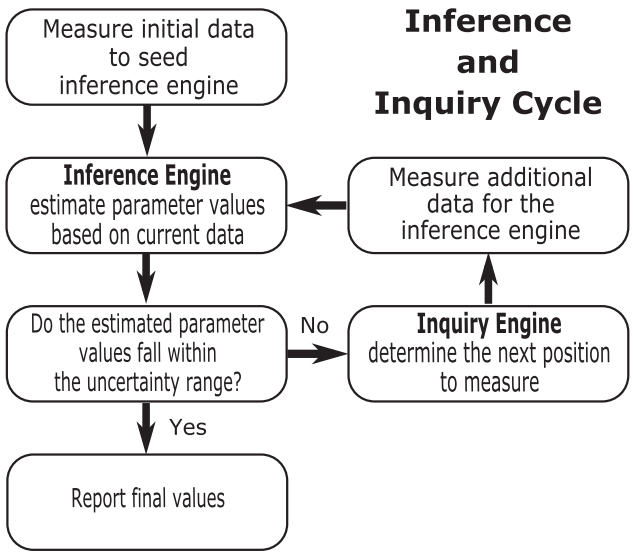

FIG. 2. Bayesian inference and inquiry cycle. The inference engine estimates the parameter values that describe the beams location and size, and the inquiry engine determines the highest quality measurement to take next.

do help to decrease the time it takes to do so. The inference engine then estimates the parameter values for the location and size of the beam utilizing the data collected so far. During this process the inference engine collects samples from the posterior probability space to aid the inquiry engine in its task of designing the next experiment or location to measure. If the estimated parameter values determined by the inference engine are within a prescribed uncertainty range then the algorithm reports the values and terminates. If the values are not within this range the samples recorded by the inference engine are passed to the inquiry engine so it can determine what location on the DMD would provide the most informative data to return to the inference engine. The algorithm then measures the light intensity of the locations determined by the inquiry engine and passes that data back to the inference engine. This cycle will continue until the parameter values that are estimated by the inference engine fall within the uncertainty range prescribed by the user.

1. Inference engine

There are number of excellent textbooks and review papers that provide a wealth of information regarding Bayesian probability theory and data analysis. These include, but are not limited to, Refs. 18–25. For this experiment, the interested reader should look specifically to Refs. 26 and 18 and especially Refs. 27 and 28 for additional information.

Bayes' theorem is applied to the beam model M with its three model parameter values $\theta_M = \{(x_0, y_0), r_0\}$ given some data D. By taking the model M and the data D as the given context, one can compute the degree (probability) to which the model M and the data D imply the hypothesized model parameter values θ_M by

$$P(\theta_M|D, M, I) = P(\theta_M|M, I)\,\frac{P(D|\theta_M, M, I)}{P(D|M, I)}. \tag{4}$$

The first term on the right-hand side of Eq. (4), $P(\theta_M|M, I)$, is the *prior probability*, which quantifies what one knows about the model parameter values θ_M conditional on one's prior information I as well as the fact that a particular model M has been hypothesized. The ratio on the right is composed of two data-dependent terms: the likelihood $P(D|\theta_M, M, I)$ and the evidence $P(D|M, I)$. The likelihood represents the degree to which the model, its hypothesized model parameter values, and the prior information could have resulted in the recorded data. As will be described in more detail, the likelihood relies on the forward (or predictive) model, which, given the hypothesized model, predicts what one would expect to observe in the course of an experiment. By comparing the predicted results to the observed results (manifested as the data D) one can quantify the probability that a physical situation described by the model could have given rise to the recorded data. The data-dependent term in the denominator is called the *evidence* or the *marginal likelihood*. In parameter estimation problems, one can conceive of this term as a normalization factor. But it can also be shown[23] that the evidence quantifies the degree to which the data D supports a given model M. This data-dependent ratio of the likelihood over the evidence modifies the prior probability for the model parameter values resulting in the *posterior probability* $P(\theta_M|D, M, I)$, which quantifies the degree to which the model, the data and the prior information together imply the hypothesized model parameter values. In this way, Bayes' theorem acts as a learning rule, which updates one's knowledge about a model after the consideration of additional data.

The learning process must be jump-started by the assignment of prior probabilities for the model parameter values. As will be described, the model parameters were assigned based on uniform probability distributions so that the analysis is dominated by the data-dependent likelihood.

As described earlier, the beam's location on the DMD is modeled as a circle with three parameters [the center position (x_0, y_0) and the radius r_0]:

$$\theta_M = \{(x_0, y_0), r_0\}\,. \tag{5}$$

The inference engine is designed to estimate these three beam parameter values using data collected from previous experiments and provide the inquiry engine with a set of posterior samples from which it can determine the next optimal experiment. This is accomplished by working with a forward model M_C that describes the basic configuration of a circle (beam), and any data d_i collected with a power

meter during the N previous light measurements made by flipping the micro mirrors at points (x_i, y_i) on the DMD.

The model utilized by the inference engine is based on the assumption that a mirror within the beams radius will produce a bright power reading if turned on, and a mirror outside of the beams radius will produce a dark power reading when turned on. At the beginning of the experiment the algorithm measures the dark field power reading and assigns a bright reading value that considers the instruments noise level. When taking a power measurement, the algorithm assigns a value of fifty to the reading if it is above the threshold, and assigns a value of twenty five for measurements below the threshold. Although the power measurements themselves have units of (W/m^2) the values assigned by the algorithm have arbitrary units. The collective measurements are labeled as

$$D = \{\{(x_1, y_1), d_1\}, \{(x_2, y_2), d_2\}, \ldots, \{(x_N, y_N), d_N\}\}. \tag{6}$$

Since the DMD consists of a discrete set of micro mirrors that can be precisely accessed, the positions of the measurements (meaning the positions of the mirrors that are flipped) are assumed to be known with certainty.

The prior probability $P(\theta_M|M_C, I)$ for the circle's parameter values were assumed to be uniformly distributed within the confines of the DMD (1024 X 768) and have a radius uniformly distributed between twenty and three hundred DMD mirrors. These assumptions resulted in the following normalized prior probability assignments

$$P(x_0|M_C, I) = (x_{max} - x_{min})^{-1} \tag{7}$$

$$P(y_0|M_C, I) = (y_{max} - y_{min})^{-1} \tag{8}$$

$$P(r_0|M_C, I) = (r_{max} - r_{min})^{-1}, \tag{9}$$

where $x_{max} = 1024$, $x_{min} = 0$, $y_{max} = 768$, $y_{min} = 0$, $r_{max} = 300$, and $r_{min} = 20$.

In the context of an optics experiment it is assumed that the entire beam will be situated on the DMD surface, however the algorithm only assumes that the center of the beam is confined to the surface of the DMD and the radius of the beam is not constrained other than a maximum and minimum size. This freedom allows for the possibility of the beam to be partially outside of the DMD's surface area, despite this freedom the algorithm would still be able to characterize the beam as long as reasonable portion of the beam remained within the boundaries of the DMD. This relaxation of the radius constraints not only increases the speed at which the algorithm operates, it also allows the algorithm the ability to explore more of the probability density.

The probability density explored by nested sampling within the inference engine is also constrained by the likelihood $P(D|\theta_M, M_C, I)$. The likelihood function quantifies the degree to which the recorded intensities could have been generated by the model using a particular set of model parameter values. A Gaussian likelihood function was used in this experiment. The general form of this function for the discretized version used for this experiment is

$$P(D|\theta_M, M_C, I) = (2\pi\sigma^2)^{-N/2} \exp\left[-\left(\frac{1}{2\sigma^2}\sum_{i=1}^{N}(M_C(x_i, y_i; x_0, y_0, r_0) - d_i)^2\right)\right] \tag{10}$$

where the forward model, M_C, *predicts* the greatest intensity of 50 units if the micro mirror location (x_i, y_i) is inside the hypothesized beam footprint at (x_0, y_0) with radius r_0, and predicts the lowest intensity of 25 units if the micro mirror location is outside of the hypothesized footprint:

$$M_C(x_i, y_i; x_0, y_0, r_0) = \begin{cases} 50, & \text{if } (x_i - x_0)^2 + (y_i - y_0)^2 \le r_0^2 \\ 25, & \text{if } (x_i - x_0)^2 + (y_i - y_0)^2 > r_0^2. \end{cases} \tag{11}$$

The simplest explanation of the likelihood function used here is, if the measured position is within the hypothesized beam footprint then it expects that the measurement will be bright, and if it is outside the hypothesized footprint it expects it to be dark. In this way the likelihood function is a function of the intensity predicted by the model. Furthermore, the parameter σ in Eq. (10) represents the expected squared deviation of the N number of measured values from the ideal values predicted by the model M_C in Eq. (11) and for the purposes of this experiment had a value of

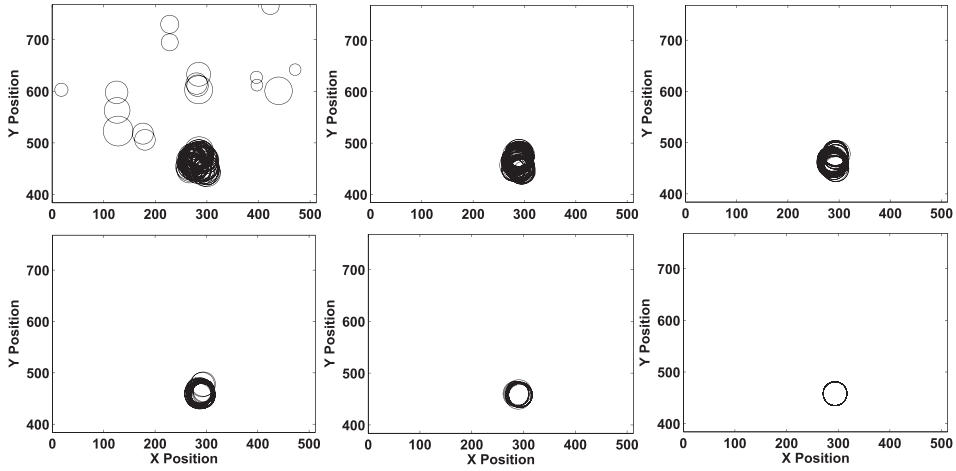

FIG. 3. The samples, (or walkers) positions within the posterior probability space are determined by their individual parameter values and thus are represented by circles. During the iterative process of nested sampling the samples move towards the most likely set of parameter values to have produced the recorded data. The series of images shown represent the progress of the samples during the final inference cycle using all of the recorded data.

five. The posterior probability Eq. (4) is proportional to the product of the prior probability and the likelihood.

Nested sampling[18] explores the posterior probability by creating and evolving a group of samples, also called walkers Fig. 3, each of which is defined by its coordinates in the model parameter space $[(x_0, y_0), r_0]$ so that it represents a single hypothesis. Nested sampling is designed to numerically integrate the posterior probability via stochastic integration. However, it also allows one to compute mean parameter values and their associated uncertainties from the posterior samples.

A more detailed explanation of nested sampling begins with the prior function creating the one hundred individual samples that will explore the posterior probability space. Each sample represented as circles in Fig. 3 are defined by the individual parameter values $[(x_0, y_0), r_0]$ that are randomly selected from within the prior probability density Eqs. (7)–(9). The likelihood associated with each sample is computed from Eq. (10) using the N number of data points recorded at that point. This value quantifies the likelihood that the parameter values represented by the individual sample could have produced the recorded data. The algorithm then determines the object that is the least likely to have produced the data and looks to replace its parameter values with values that have a higher likelihood. It chooses these new values by randomly exploring the volume around another more likely sample and moving the sample in question to the new location. This procedure is repeated in an iterative process that moves the samples inexorably towards the most likely parameter values as shown in Fig. 3 (bottom right). During the process the algorithm stores the discarded sample values and uses them to perform the integration as well as the mean model parameter values and their uncertainties by taking weighted averages.[18]

Given the saved sample values generated by the nested sampling algorithm, a set of samples can be generated from the posterior. These posterior samples, which provide a representative set of possible model parameter values, are passed to the inquiry engine, which uses them to determine the optimal location at which to take the next measurement.

2. *Inquiry engine*

The inquiry engine, armed with the posterior samples provided by the inference engine, looks to determine the next position on the DMD to measure, or equivalently, the optimal experiment to perform. Drawing on the theory provided by Ref. 26 one can consider the proposed experiment E as taking a measurement at position (x_e, y_e). Although the measurement value or the circle's true parameter values are not known, the probability of the measured intensity d_e in terms of the joint probability of d_e and θ_M can be written as

$$P\left(d_e|D,\left(x_e,y_e\right),I\right)=\int P\left(d_e,\theta_M|D,\left(x_e,y_e\right),I\right)\mathrm{d}\theta_M. \tag{12}$$

Employing the product rule Eq. (12) can be written as

$$P\left(d_e|D,\left(x_e,y_e\right),I\right)=\int P\left(d_e|\theta_M,D,\left(x_e,y_e\right),I\right)P\left(\theta_M|D,\left(x_e,y_e\right),I\right)\mathrm{d}\theta_M. \tag{13}$$

This can be simplified by noting that if the model parameter values θ_M were known then the data D would not be needed, so the dependence of the first probability under the integral for d_e on D can be omitted:

$$P\left(d_e|D,\left(x_e,y_e\right),I\right)=\int P\left(d_e|\theta_M,\left(x_e,y_e\right),I\right)P\left(\theta_M|D,\left(x_e,y_e\right),I\right)\mathrm{d}\theta_M. \tag{14}$$

Probability theory does not allow one to make a precise decision. To decide on an optimal measurement location $(\hat{x}_e,\hat{y}_e)$, one needs to define a utility function U(outcome, action) which can be optimized with respect to the probability of the expected measured intensity d_e

$$\left(\hat{x}_e,\hat{y}_e\right)=\int P\left(d_e|D,\left(x_e,y_e\right),I\right)U\left(d_e,\left(x_e,y_e\right)\right)\mathrm{d}d_e. \tag{15}$$

where the measurement location (x_e,y_e) represents the action and the measurement result d_e represents the outcome.[27] Since the aim is to take measurements that maximize the expected information gain, the Shannon information was employed as the utility function

$$U\left(d_e,\left(x_e,y_e\right)\right)=\int P\left(\theta_M|d_e,D,\left(x_e,y_e\right),I\right)\log P\left(\theta_M|d_e,D,\left(x_e,y_e\right),I\right)\mathrm{d}\theta_M. \tag{16}$$

By writing the integrals for the joint entropy for θ_M and d_e two ways, it can be shown[26] that the experiment that promises to provide the most information is the one that maximizes the entropy of the distribution of possible measurements:

$$\left(\hat{x}_e,\hat{y}_e\right)\equiv argmax_{(x_e,y_e)}\left(-\int P\left(d_e|D,\left(x_e,y_e\right),I\right)\log P\left(d_e|D,\left(x_e,y_e\right),I\right)\mathrm{d}d_e\right). \tag{17}$$

That is, the best place to measure is the place where you are least certain as to what the result will be.

The inquiry engine works by considering each possible measurement location (x_e, y_e) and uses the set of posterior samples provided by the inference engine along with the likelihood Eq. (10) to make a set of predictions of what d_e could be measured. This set of predictions of d_e for each measurement location represents a set of samples from the probability density of d_e. From this, the entropy of the density of d_e can be estimated by constructing an optimal histogram model of the density function from the samples,[29] and the experiment $E=(\hat{x}_e,\hat{y}_e)$ with the greatest entropy is selected for the next experiment, which will in turn provide additional data for the inference engine completing the inference-inquiry learning loop. This inference-inquiry process is repeated until the model parameters are estimated within a pre-defined tolerance.

II. EXPERIMENT

A. Experimental design

The optical setup Fig. 4 consisted of a linearly polarized Research Electro-Optics Inc. model 30995 HeNe laser that operated at 17 mW with a wavelength of 633 nm. The laser beam passed through a spatial filter consisting of microscope objective and a five μm pinhole aperture. The beam then passed through an iris and a f=50 mm parabolic lens collimating the beam. The beam passed

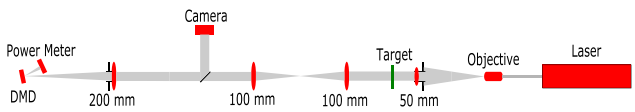

FIG. 4. The optical setup with the beam entering from the right and focusing on the DMD in the image Fourier plane.

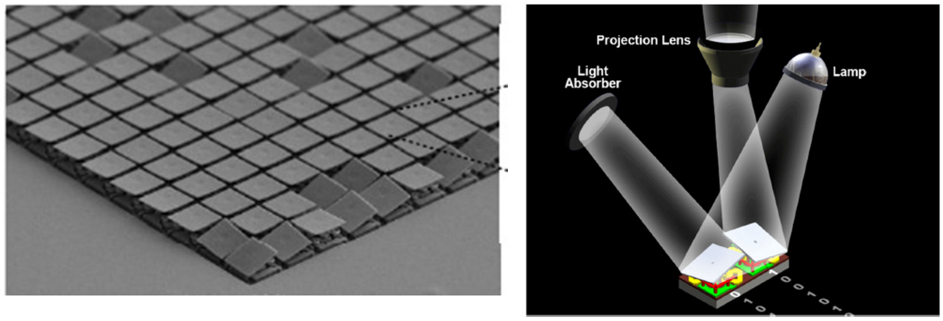

FIG. 5. A DMD array of micro mirrors (a) and their function (b) (Courtesy of Texas Instruments).

through the remaining lenses with the DMD placed at the Fourier plane. The DMD used for the experiment was manufactured by Texas Instruments Incorporated and had a resolution of 1024 X 768 mirrors and was controlled by the Vialux ALP-4.1 Control Suite by Pixel Precision.

The DMD itself consists of an array of 13.68 micron micro mirrors Fig. 5(a) that can be individually addressed and operated through Matlab. The mirrors exist in a state of either tilting twelve degrees in the positive or negative direction Fig. 5(b). In the optical setup Fig. 4 the DMD was held at 12 degrees in relation to the optical axis allowing it to either reflect the beam back through a beam splitter to produce an image at the Point Gray camera or in the other direction to the Thorlabs digital power meter. This selection could be accomplished at the pixel level creating a very precise binary filter at the Fourier plane.

The algorithm begins by directing all of the mirrors on the DMD away from the power meter and then towards it to establish a baseline dark and full light reading. Armed with these values the algorithm divides the DMD into four sections to facilitate a coarse search for the beam's location. It does this by directing all of the mirrors in a particular quadrant towards the power meter and eliminates quadrants that do not produce a reading above the dark baseline. This allows the algorithm to quickly narrow the playing field down considerably. This process could be continued but it was decided to place the burden of locating the circle on the algorithm.

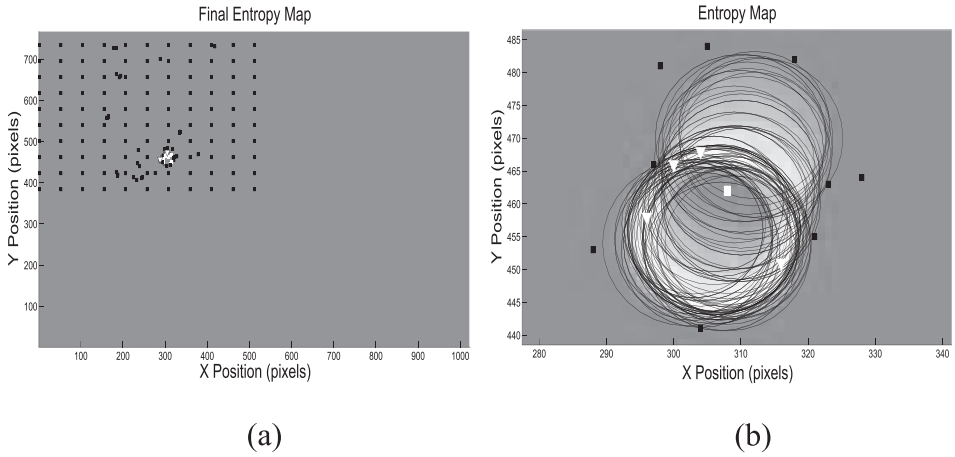

FIG. 6. The inquiry engine produces an entropy map that displays the regions in the posterior probability space that would produce the most informative measurement results shown as a lighter region of gray (a). A closeup view of the region of interest (b) shows points that have been tested but were dark (black squares), points that were light (white squares), and the four points that would yield the most information in the search for the beam (white triangles). These points would then be measured and added to the total data set and passed back to the inference engine.

TABLE I. Final Parameter Values.

Parameter	Value	Uncertainty
x_0 (mirrors)	304	.001
y_0 (mirrors)	457	.002
r_0 (mirrors)	16	.01

The algorithm then selects between one and four testing points that are uniformly distributed within the remaining playing field, this is done to seed the inference engine with initial data further reducing the overall time required for the task. The inference engine then uses this data to estimate the beam's size and location while exploring the posterior probability density. In addition to providing the estimated parameter values with their uncertainties, the inference engine provides samples drawn from the posterior probability density that will be used by the inquiry engine to determine the next experiment.

The inquiry engine then uses the samples provided by the inference engine to create a map of the expected information gain at each point in the playing field illustrated graphically in Fig. 6. The lighter shades of gray represent higher regions of entropy and thus points sampled from within these regions will provide data with the highest probability of information gain. The algorithm then selects several points from the highest entropy region represented as triangles in Fig. 6 and measures their intensity values. It then adds the new data to the data set and passes it all to the inference engine for consideration. This process of inference and inquiry continues until the data collected provides an estimation of the beam's size and position with an uncertainty that is below a specified precision. When the specified precision is reached the algorithm reports its final results and tests the validity of the chosen parameter values.

B. Experimental results

The algorithm was run with the system configured as illustrated in Fig. 4 without the target. This allowed only the collimated illuminating beam to be considered. The collimated beam travels nearly parallel to the optical axis creating a very concentrated point of light at the center of the Fourier

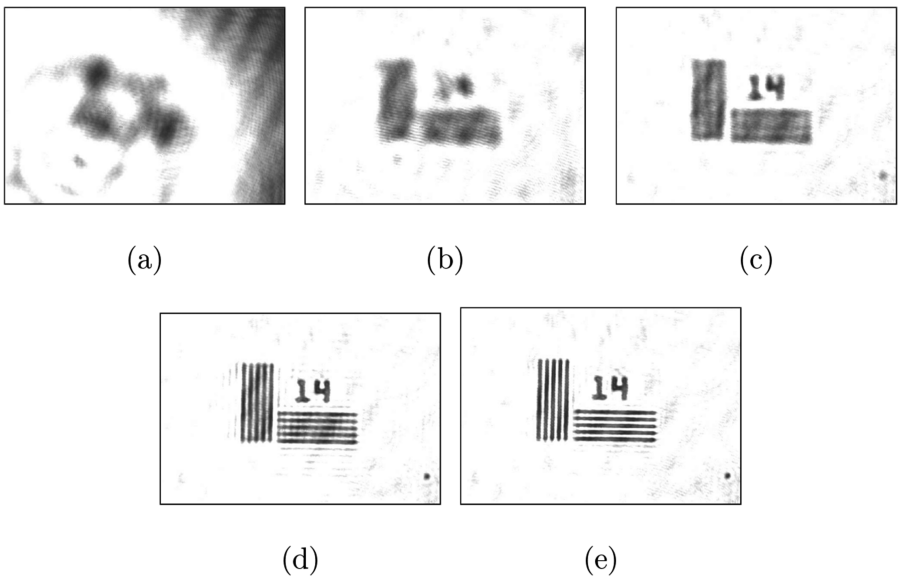

FIG. 7. Shows the resulting image of a high resolution image of the Thorlabs target when the DMD creates a circle of mirrors that starts at r=2 mirrors (a) and then increases to r=30 mirrors (b), then to r=50 mirrors (c), then to r=75 mirrors (d) then to r=100 mirrors (e).

plane. Although the beam of light is most accurately modeled as a wide Gaussian beam passing through a circular aperture, it is accurate enough for the purposes of this experiment to model it as a circle of uniform intensity. The central region of the beam propagating through the system is nearly circular, and the utilization of a more complicated model would not have provided significantly superior parameter estimations.

The experiment produced the final parameter estimates listed in Table I. The search required fifteen cycles of inference and inquiry producing a total of one hundred and sixty nine data points; this number includes the initial seed data distributed throughout the 512 x 384 mirror playing field.

In an effort to verify the size and position of the beam at the Fourier plane a series of tests were performed. The system first directed a circle of mirrors determined by the inference engine to be the beam's size and location towards the power meter. As expected the power meter recorded a power level equal to the initial full light reading recorded at the beginning of the experiment. It then directed those same mirrors away from the power meter and the remaining mirrors toward the meter.

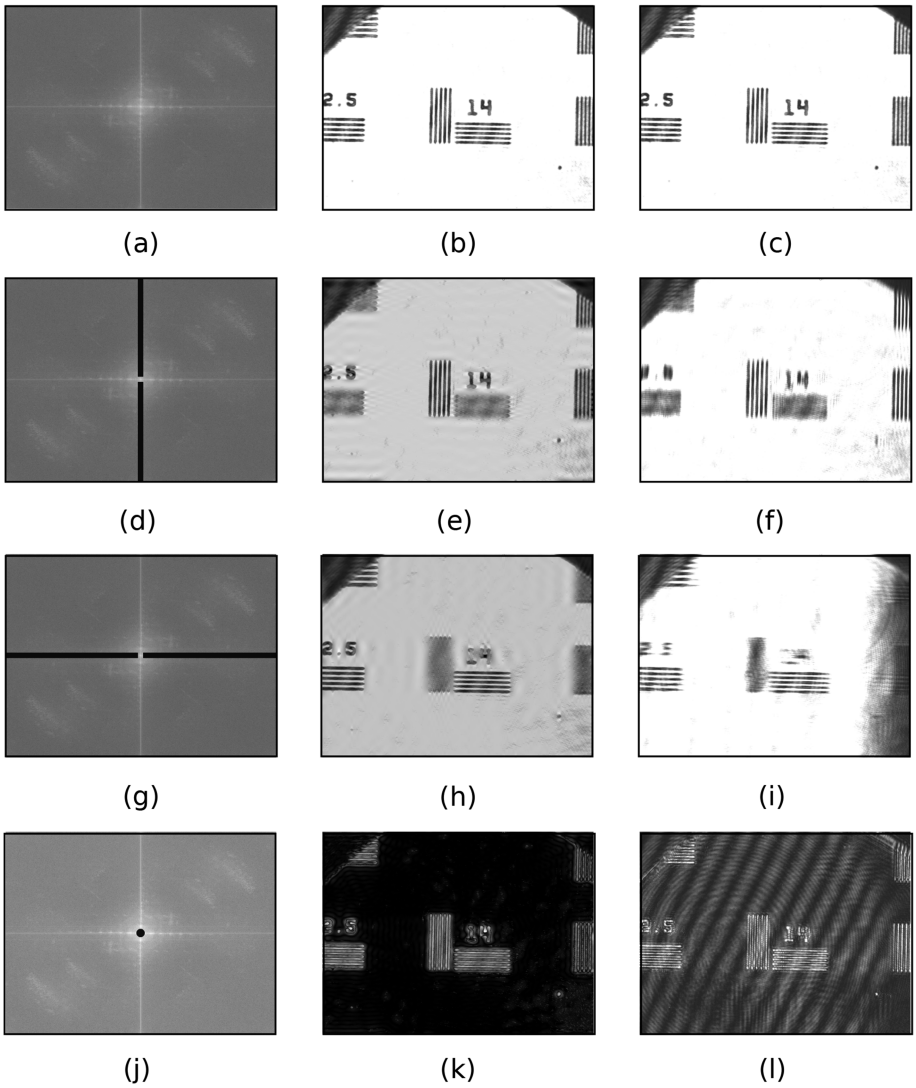

FIG. 8. An image of the Thorlabs target was captured by the camera (b) and (c) and a Fourier transformation was computationally performed with Matlab (a). A Fourier plane mask was then applied (d), (g) and (j), both computationally (center column) and physically with the DMD (right column). It can be seen that the resulting images are remarkably similar which proves that the DMD is acting as a functional Fourier filter with proper alignment.

This resulted in a power reading equal to the dark baseline reading taken at the beginning of the experiment. This initial procedure supports the claim that the algorithm had correctly determined the size and position of the illuminating beam.

In order to further verify that the center and size of the beam in the Fourier plane had been correctly characterized, an image of a Thorlabs resolution target was acquired with this system. A variety of Fourier plane filters were then applied both computationally and physically using the DMD. The parameter values used for the physical manipulation of the Fourier plane by the DMD were those provided by the inference engine.

A procedure was then performed to verify the concentricity between the determined beam location and the center of the physical Fourier plane. As described in Section 1.A., resolution depends on the separation between mirrors that are redirecting light to the camera: the resolution of features oriented perpendicular to the mirror separation improves as mirror separation increases. Therefore, a series of circles were created on the DMD at the beam's determined location with a radius that varied from r=0 to r=100 mirrors in ten mirror increments until the entire image was in focus Fig. 7. The resulting images were recorded at each iteration for analysis. If the circle of mirrors and the beam were concentric, the resolution target features in both the horizontal and vertical planes should improve at the same rate. This predicted behavior is exactly what was recorded verifying the concentricity of the beam.

In order to verify the alignment of the system, filters were used to reduce the resolution in two orthogonal directions in both simulation and experiment. The Fourier masks used, and the resulting Matlab images are displayed in the first two columns of Fig. 8. The third column of Fig. 8 is the result of the same mask displayed on the DMD at the Fourier plane. The mask shown in Fig. 8(d) with results in Figs. 8(e) and (f) is designed to produce a blurry image in the vertical direction while preserving horizontal sharpness. The mask shown in Fig. 8(g) with results in Figs. 8(h) and (i) is designed to blur in the horizontal, but not vertical direction. The mask shown in Fig. 8(j) with results in Figs. 8(k) and (l) is designed to remove the incident beam and preserve only light that has been perturbed by the sample, highlighting its edges on a dark background. There is good agreement between the simulated results and the experimental results. However there are some discrepancies that are likely due to stray reflections and diffractive effects not included in the simulation. This is particularly evident in Figs. 8(k) and (l).

III. SUMMARY

The size and location of a coherent beam of light reflecting off of a 1024 x 768 DMD which was located in the Fourier plane of a 4f lens system was estimated utilizing a Bayesian inference and inquiry algorithm. The search was facilitated by measuring the power of the light reflected off of individual mirrors of the DMD which were selected for measurement by the algorithm.

The algorithm began with a coarse binary search which narrowed the remaining area to one quarter of the entire DMD, it then measured one hundred and eleven evenly spaced points within the remaining area to obtain initial data to work with. The algorithm then performed fifteen cycles of inference and inquiry to select fifty eight points within the search area that it determined to provide the most information for the search.

These results are in sharp contrast to the expected performance of a brute force search. If one were to consider a playing field that was not much larger than the circle itself, a brute force search would require a considerable portion of the circle's approximately eight hundred mirrors to be tested in order to characterize the circle's parameters to the precision realized with the algorithm. In contrast the algorithm itself would require approximately ten cycles of inference and inquiry to accomplish the same task.

This marked improvement in performance is due to the learning capabilities of Bayes' theorem. It enables the algorithm to learn from the data collected and consider which experiment will provide the most information to further its knowledge. Although in the described experiment, this capability has been applied to finding a circle of light within a field of mirrors, the same concept can be applied to any situation that requires a system to not only perform data analysis but to also design the experiments autonomously.

[1] N. Chattrapiban, E. A. Rogers, D. Cofield, W. T. Hill III, and R. Roy, "Generation of nondiffracting Bessel beams by use of a spatial light modulator," Opt. Lett. **28**, 2183–2185 (2003).
[2] J. A. Davis, J. Guertin, and D. M. Cottrell, "Diffraction-free beams generated with programmable spatial light modulators," Appl. Opt. **32**, 6368–6370 (1993).
[3] V. Bagnoud and J. D. Zuegel, "Independent phase and amplitude control of a laser beam by use of a single-phase-only spatial light modulator," Opt. Lett. **29**, 295–297 (2004).
[4] Y.-X. Ren, M. Li, K. Huang, J.-G. Wu, H.-F. Gao, Z.-Q. Wang, and Y.-M. Li, "Experimental generation of Laguerre–Gaussian beam using digital micromirror device," Appl. Opt. **49**(10), 1838–1844 (2010).
[5] L. Gong, Y.-X. Ren, G.-S. Xue, Q.-C. Wang, J.-H. Zhou, M.-C. Zhong, Z.-Q. Wang, and Y.-M. Li, "Generation of nondiffracting Bessel beam using digital micromirror device," Appl. Opt. **52**(10), 4566–4575 (2013).
[6] M. Mirhosseini, O. S. Magaña-Loaiza, C. Chen, B. Rodenburg, M. Malik, and R. W. Boyd, "Rapid generation of light beams carrying orbital angular momentum," Opt. Express **21**, 30196–30203 (2013).
[7] Y.-X. Ren, R.-D. Lu, and L. Gong, "Tailoring light with a digital micromirror device," Ann. der Phys. **527**(7-8), 447–470 (2015).
[8] B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. Padgett, "3D computational imaging with single-pixel detectors," Science **340**(6134), 844–847 (2013).
[9] M. F. Duarte *et al.*, "Single-pixel imaging via compressive sampling," IEEE Signal Processing Magazine **25**(2), 83–91 (2008).
[10] Y. Qiao, X. Xu, T. Liu, and Y. Pan, "Design of a high-numerical-aperture digital micromirror device camera with high dynamic range," Appl. Opt. **54**(1), 60–70 (2015).
[11] W. Feng, F. Zhang, W. Wang, W. Xing, and X. Qu, "Digital micromirror device camera with per-pixel coded exposure for high dynamic range imaging," Appl. Opt. **56**(13), 3831–3840 (2017).
[12] J. A. Davis, D. E. McNamara, D. M. Cottrell, and T. Sonehara, "Two-dimensional polarization encoding with a phase-only liquid-crystal spatial light modulator," Appl. Opt. **39**, 1549–1554 (2000).
[13] A. M. Weiner, "Femtosecond pulse shaping using spatial light modulators," Review of Scientific Instruments **71**(5), 1929–1960 (2000).
[14] F. Zernike, "How i discovered phase contrast," Science **121**(3141), 345–349 (1955), Bibcode:1955Sci. . . 121..345Z, PMID 13237991.
[15] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," IEEE Signal Processing Magazine **32**(3), 87–109 (2015).
[16] E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," SIAM Review **57**(2), 225–251 (2015).
[17] J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts and Company, Englewood), Ch. 6 (2005).
[18] D. S. Sivia and J. Skilling, *Data Analysis A Bayesian Tutorial*, 2nd edition (Oxford University Press, Oxford, 2010).
[19] R. T. Cox, "Probability, frequency and reasonable expectation," Am. J. Phys. **14**, 1–13.
[20] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, 2003).
[21] P. Gregory, *Bayesian logical data analysis for the physical sciences*, Vol. 10, (Cambridge University Press, Cambridge, UK, 2003).
[22] J. V. Candy, *Bayesian Signal Processing: Classical, Modern and Particle Filtering Methods* (John Wiley and Sons, Inc., 2009).
[23] K. H. Knuth, M. Habeck, N. K. Malakar, A. M. Mubeen, and B. Placek, "Bayesian evidence and model selection," Digital Signal Processing **47**, 50–67 (2015), arXiv:1411.3013 [stat.ME].
[24] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*, 3rd edition (CRC Press, 2013).
[25] W. von der Linden, V. Dose, and U. von Toussaint, *Bayesian Probability Theory: Applications in the Physical Sciences* (Cambridge University Press, 2014).
[26] T. J. Loredo 2003, Bayesian adaptive exploration, In: G. J. Erickson, Y. Zhai (eds) Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Jackson Hole WY, USA, AIP Conf. Proc. **707**, Melville NY: AIP, pp. 330–346.
[27] K. H. Knuth, P. M. Erner, and S. Frasso, "Designing intelligent instruments," in Proceedings of the 27th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science ans Engineering, K. Knuth, A. Caticha, J. Center, A. Giffin and C. Rodriguez, Eds., vol. 954, pp. 203–211, AIP, July 2007.
[28] N. K. Malakar, D. Gladkov, and K. H. Knuth, "Modeling a sensor to improve its efficacy," Journal of Sensors **2013**, 11 (2013), Article ID 481054, arXiv:1303.4385 [physics.ins-det].
[29] K. H. Knuth, "Optimal data-based binning for histograms," (arXiv:physics/0605197v2).