Optimal Frequency Sampling for Detection in Multi-Spectral Images

Leonhard Grosse, *Student Member, IEEE*, Markus Flierl, *Member, IEEE,*

***Abstract*—Substance detection using low-resolution multi-spectral imaging has been proven to be useful for use-cases in intelligent farming and disease control. Adapted frequency sampling strategies are known to improve detection performance when scenario-specific reflection characteristics of materials are used. In this paper, we propose a framework for the evaluation and construction of optimal frequency sampling schemes under the assumption of a linear mixing model using concepts from optimal experiment design. We then motivate D-optimal designs using an analytically traceable example. Consequently, we derive a set of Greedy algorithms with lower computational complexity that approximate such D-optimal designs. We provide numerical examples that support the proposed sampling schemes and show that our framework is beneficial for substance detection and abundance estimation for controlling mosquito larvae in water.**

***Index Terms*—Multi-spectral imaging, linear mixing model, experiment design, Fisher information, greedy algorithms.**

I. Introduction

THE availability of low-cost multi-spectral sensors has resulted in the exploration of a variety of use-cases for multi-spectral imaging. The increased amount of information multi-spectral sensors are able to collect from visual scenery has been proven to be especially useful for applications in the agricultural sciences [1], [2], [3]. More recently, various publications explore the use of unmanned aerial vehicles (UAVs) equipped with multi-spectral cameras for efficient disease control in regions with problematic populations of disease-bearing mosquitoes [4], [5], [6]. These techniques are largely dependent on efficient classification of ground material using the noisy multi-spectral images obtained by the system. This closely connects to the field of remote sensing, which typically uses high-resolution hyper-spectral images obtained from satellites for environmental monitoring. The techniques used for substance identification presented, e.g., in [7], [8], give a promising outlook on further developing low-resolution multi-spectral sensor-based techniques that enable more precise detection of larval habitats, and therefore increase efficiency of health authorities in disease prevention. While high-resolution sensors sample the continuous reflectance spectra so detailed that the samples allow near-optimal reconstruction, low-resolution multi-spectral sensors only capture selected frequency positions, while disregarding large areas of the spectral characteristics that - if sampled - might significantly increase detection performance. Formalisms that are able to distinguish the quality of sampling strategies, i.e., different choices of sampling frequencies, as well as algorithms to find detection-optimal sampling positions are therefore desirable.

L. Grosse and M. Flierl are with KTH Royal Institute of Technology in Stockholm, email: {lgrosse, mflierl}@kth.se.

For general sensor placement problems, as well as selected use-cases from other specific problem areas like electrical grid control, there exists a substantial body of research that considers strategies from optimal experiment design to tackle efficient selection or placement (see , e.g., [9], [10], [11]).

A. Main Contributions

In this paper, we formulate a context-specific mathematical framework for sampling frequency selection and evaluation in the multi-spectral imaging for substance detection context, utilizing estimation-theoretic arguments and tools from optimal experiment design. Using an analytically traceable example, we show that in order to be in line with intuition, D-optimal designs are to be preferred over, e.g., A-optimal designs in this context. Based on the derived model, we derive a Greedy-algorithm that is able to approximate D-optimal designs for frequency sampling in linear time. Finally, by simulating detection with numerical data available from real hyper-spectral measurements, we give a first example that hints at the advantage of problem-specific sampling frequency placement for substance detection and demonstrate how the presented techniques can be used in a well-defined use-case.

B. Related Work

Multi-spectral imaging has received attention in various research areas. This work focuses, and is in parts based on, the information-theoretic analysis of hyper-spectral imaging systems in fluorescence microscopy presented by Ram [12], in which the author formulates a stochastic model for systems using an arbitrary number of sampling frequencies (or channels), and - using the notion of Fisher Information - derives performance bounds on abundance estimation under the consideration of three different stochastic data models. In order to extend the framework constructed in this way to sampling frequency selection, we explore the field of optimal experiment design, which, e.g., the authors of [11] have shown to be useful in electrical grid management. Significant and extensive works on optimal experiment design include the works of Pázman [13] and Patan [9], with the latter being more recent, considering a highly general sensor placement problem based on partial differential equations in continuous time and space. Ucinski and Patan [14] further present an approach for solving a discrete sensor placement problem using a depth-first branch-and-bound method, which they prove to be effective in avoiding sensor clustering, overcoming this common difficulty with the continuous formulation. Other algorithmic works on optimal design include the works on sensor selection by Nagata *et al.* [10] and Saito *et al.* [15], which present

and test algorithms for A-optimal and D-optimal designs, respectively. The authors of the latter present a Greedy-strategy for optimization that differs from the one presented in this work due to a largely different mathematical formulation of the underlying problem.

C. Outline of the Paper

We start by introducing our theoretical model, alongside the corresponding notation in Section II. Then, we restate some central results and concepts from parameter estimation and experiment design in Section III. We present the illustrative and motivating example in Section IV. In Section V, we present algorithmic solutions. Section VI presents all numerical results. Section VII concludes the paper.

II. Spectral Unmixing and the Linear Mixing Model

We consider a set of S substances $\mathbb{S} = \{s_i \mid i = 1, \dots, S\}$ to distinguish between, where each substance emits a characteristic reflectance spectrum $\Phi_i(f)$. Given a fixed budget of N sampling frequencies, we aim to determine the optimal sampling positions $\boldsymbol{f} \equiv (f_1, \dots, f_N)^T$, such that detection of a substance mixture is possible with the lowest possible error in the mean squared error sense. As shown in, e.g., [8], for the case of straight-forward detection without any specific assumptions on optical interactions between the materials, the physical reflection is well approximated by a *linear mixing model* (LMM) of substance-dependent reflectances, i.e., the spectral reflectance $\Phi^p(f)$ of ground material at any specific pixel p is given by

$$\Phi^p(f) = \theta_1 \Phi_1^p(f) + \dots + \theta_S \Phi_S^p(f),$$

with the natural constraints on the mixture coefficients, also refered to as *abundaces* θ_i,

$$0 \leq \theta_i \leq 1 \quad \forall i, \qquad \sum_{i=1}^{S} \theta_i = 1. \tag{1}$$

Since all of the following derivations can be done by-pixel without taking spatial features into account, the index p will be omitted for compact notation.

Making use of the constraints in (1), the LMM results in a linear mapping from the $(S-1)$-dimensional coefficient space to a vector $\boldsymbol{\Phi}(\boldsymbol{f}) = (\Phi(f_1), \dots, \Phi(f_N))^T \equiv \boldsymbol{A}\boldsymbol{\theta}$ in the N-dimensional sampling space, expressed using a mixing matrix $\boldsymbol{A} \in \mathbb{R}^{(S-1)\times N}$ and an abundance-vector $\boldsymbol{\theta} \in \mathbb{R}^{(S-1)}$ as

$$\boldsymbol{A}\boldsymbol{\theta} \equiv \begin{pmatrix} D_1(f_1) & \dots & D_{S-1}(f_1) \\ \vdots & \ddots & \vdots \\ D_1(f_N) & \dots & D_{S-1}(f_N) \end{pmatrix} \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_{S-1} \end{pmatrix},$$

with $D_i(f_n)$ indicating the difference of the spectrum Φ_i to the spectrum Φ_S at sampling frequency f_n

$$D_i(f_n) \equiv \Phi_i(f_n) - \Phi_S(f_n).$$

Following the discussions on multi-spectral data models in [12], we model the sensor observations of the resulting mixture spectrum at the sampling frequencies as being perturbed by white Gaussian noise with frequency-dependent variance, i.e., given a (deterministic) reflectance spectrum $\boldsymbol{\Phi}(\boldsymbol{f})$, the sensor observations $\boldsymbol{x}$ are assumed to be distributed according to

$$\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \Phi(f_1) \\ \vdots \\ \Phi(f_N) \end{pmatrix}, \begin{pmatrix} \sigma(f_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma(f_N) \end{pmatrix} \right). \tag{2}$$

We get the *spectral unmixing problem* as the overall parameter estimation problem

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{x}|\boldsymbol{\theta}),$$

where $p(\boldsymbol{x}|\boldsymbol{\theta})$ denotes the probability density function of the observations defined by (2). Apart from the interpretation of the characteristics as a mixture of different substances as given in the LMM, the abundance vector $\boldsymbol{\theta}$ has another interpretation: Assume that the true, underlying parameter vectors only have one non-zero element equal to one. Due to the constraints in (1), the parameter vector estimated according to (II) under this assumption can be interpreted as the estimated probabilities of each substance being present in the pixel under consideration. While the mixture interpretation resembles the common abundance estimation problem known from remote sensing, this interpretation is more closely connected to classification as it appears in neural network literature, in which the output of a classifier - if well calibrated - has an interpretation as the posterior probabilities of the input belonging to each class, assuming deterministic class membership [16], [17]. We refer to this interpretation as the *singular case*, and compare correct detection probabilities in this case in Section VI.

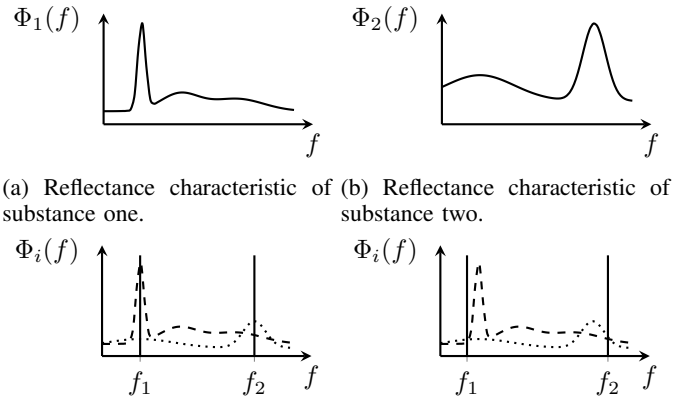
Figure 1 shows an example setup with the goal of estimating two abundances given the spectral characteristics depicted in Subfigure 1a and 1b. Subfigure 1c and 1d illustrate how even relatively small differences in the choice of the two sampling frequencies in this example lead to stark differences in estimation accuracy: While for Subfigure 1c, the two spectra differ largely in their respective reflection values at both sampling frequencies, for Subfigure 1d the two spectra are close to identical at the chosen frequencies. This will naturally make estimation less accurate under noisy measurements, with the extreme case of identical values at both sampling points, where reliable estimation becomes impossible even without any additional sensor noise. The following section is aimed at quantifying this intuition.

III. Fisher Information and Experiment Design

We will introduce the main concepts of experiment design as needed for this work in the following. We start by defining the core measure of information used in this paper, that is, the Fisher information.

A. The Fisher Information Matrix

For parameter estimation problems of the form (II), the Fisher Information is a standard tool for measuring how well observations of a parameterized random variable determine a set of these parameters and is commonly used in both



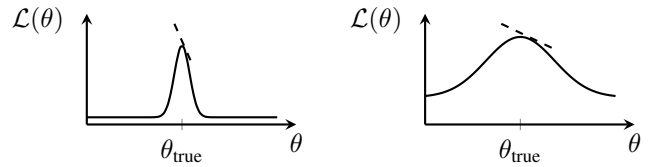
(a) Reflectance characteristic of substance one.

(b) Reflectance characteristic of substance two.

(c) Well placed sampling frequencies for discrimination between substance one and two.

(d) Poorly placed sampling frequencies for discrimination between substance one and two.

Fig. 1: Example for a two-spectrum discrimination setup and two choices of sampling frequencies.



(a) Larger slope, larger Fisher Information

(b) Smaller slope, smaller Fisher Information

Fig. 2: Comparison of two likelihoods that illustrate how the Fisher Information is connected to the curvature of the log-likelihood functions around the true parameter value

estimation and classification theory [18], [19]. It is closely related to the information-theoretic relative entropy measure, as it can be interpreted as quantifying the curvature of a conditional distribution with respect to its parameters [19], [20]. For a single parameter, i.e., in the case of $\boldsymbol{\theta} = \theta \in \mathbb{R}$, the *Fisher Information* (FI) is defined as the curvature of the log-likelihood $\mathcal{L}(\theta) = \log\, p(x|\theta)$ of the parameter θ, given an observation of a random variable parameterized by θ, that is,

$$I(\theta) = -\mathbb{E}\left[\frac{\partial^2 \log p(x|\theta)}{\partial \theta^2}\right],$$

assuming that certain regularity conditions on the parameterized family of distributions $p(x|\theta)$ hold [21]. Figure 2 illustrates this connection between higher FI and observations that more sharply define a true underlying parameter value θ_{true}, as the increased information content about a parameter of an observation results in a likelihood function for which the main mass is more concentrated around that true parameter value. For the present multi-parameter case, the elements of the *Fisher Information Matrix* (FIM) are defined as

$$(\boldsymbol{I}(\boldsymbol{\theta}))_{i,j} = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log p(x|\boldsymbol{\theta}) \middle| \boldsymbol{\theta}\right],$$

where now - in addition to the diagonal elements, which provide the single-parameter Fisher Information of each of the parameters - the non-diagonal elements provide a measure of how the estimation accuracy for the different parameters in the parameter vector each pairwise influence each other. Following the derivations made in [12] for our assumed linear mixing model and the Gaussian additive noise assumption, we obtain the elements of the FIM for (2) as

$$\begin{aligned}\boldsymbol{I}(\boldsymbol{f}) &= \boldsymbol{A}^T \text{diag}\left[\frac{1}{\sigma(f_1)}, ..., \frac{1}{\sigma(f_N)}\right]\boldsymbol{A} \\ &= \left(\sum_{n=1}^{N} \frac{D_i(f_n) D_j(f_n)}{\sigma(f_n)}\right)_{i,j},\end{aligned} \tag{3}$$

with $i, j = 1, \ldots, S-1$ and consequently $\boldsymbol{I}(\boldsymbol{f}) \in \mathbb{R}^{(S-1)\times(S-1)}$. Note that, for emphasizing the FIMs dependence on the chosen sampling frequencies, its argument has been changed to be the sampling frequency vector $\boldsymbol{f}$. We will use these two notations equivalently in the following. Equation (3) again uses the spectral difference function defined in (II). It is easily seen from this that the diagonal elements of the matrix (3) constitute the squared spectral difference at each single sampling frequency, scaled by the noise variance at this frequency. The off-diagonal elements provide information about the inter-sampling point relationship of this distance function. We will more closely examine this relationship in Section IV.

B. Optimality Measures and Optimal Experiment Design

The single parameter FI is defined on the real numbers, and therefore its extremization is conceptually straightforward. However, in the multi-parameter case, the problem of defining a matrix measure that defines optimality in a way that is suitable for a specific problem arises. To this end, different measures of a FIM are discussed extensively in the optimal experiment design literature, and there exists a multitude of different measure propositions and standards with associated interpretations [9], [13]. In general, we want to define a measure $\Psi(\boldsymbol{I}(\boldsymbol{f})) \in \mathbb{R}$ that is suitable for optimization over the set of sampling frequencies $\boldsymbol{f}$. More specifically, we want a maximization/minimization of this measure to yield sampling frequencies that follow the reasoning of spectral difference discussed in the previous section.

Many of the common measures used in optimal experiment design are connected to the uncertainty ellipsoid E_ϵ corresponding to the parameter estimation problem, which provides a region around the true parameter value, in which an estimation from the associated observations will lay with a certain probability associated with the threshold parameter ϵ [9]. Its connection to the Fisher Information Matrix is obvious from its definition:

$$E_\epsilon(\hat{\boldsymbol{\theta}}) = \{\boldsymbol{\theta} : (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T \boldsymbol{I}(\boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq \epsilon\}.$$

Therefore, for minimizing the region of uncertainty, we strive to minimize the volume of the ellipsoid E_ϵ for all possible values of ϵ. From a probabilistic viewpoint, this can be seen as making the parameter set as well-determined by any observation as possible. By noting that $\text{Vol}(E_\epsilon) \propto \det(\boldsymbol{I}(\boldsymbol{f})^{-1}) \propto 1/\det(\boldsymbol{I}(\boldsymbol{\theta}))$, it becomes clear that maximization of the FIMs determinant will be equivalent to this minimization of the

estimation uncertainty. This measure, together with the maximization task is referred to in literature as a D-optimal experiment design. While other measures are frequently used, e.g., A-optimality with $\Psi(\boldsymbol{I}(\boldsymbol{f})) = \text{tr}(\boldsymbol{I}(\boldsymbol{f}))$ (this corresponds to minimizing the ellipsoids average axis length), the geometrical considerations above suggest D-optimality to be the most holistic measure with respect to minimizing estimation uncertainty. This notion will be further motivated in Section IV, where we show that the results that are aligned with basic intuition in a simple, analytically traceable example setup can not be generated by, e.g., A-optimal designs.

IV. An Illustrative First Example: Spectral Indicator Functions

To motivate the choice of optimality criterion, as well as to gain an intuition about different sampling strategies and their connection to substance estimation in our model, it is useful to take a look at a simple example scenario where intuition for sampling strategies is immediately evident: Consider three substances S_1, S_2, S_3, for which the spectra are each given by the non-overlapping indicator functions $I_1(f), I_2(f), I_3(f)$:

$$I_i(f) \equiv I_{[c_i-\Delta, c_i+\Delta]}(f) \quad c_i+\Delta < c_{i+1}-\Delta, \quad i = 1, 2, 3.$$

Further, assume the noise variance at the sampling frequencies to be constant as $\sigma(f) = 1$. Substituting (IV) into (3), we get the Fisher Information Matrix (**??**). From this, we can derive the D-optimality criterion for this specific case as

$$\begin{aligned}\det(\boldsymbol{I}(\boldsymbol{f})) = &\sum_n \frac{(I_1(f_n) - I_3(f_n))^2}{\sigma(f_n)} \sum_n \frac{(I_2(f_n) - I_3(f_n))^2}{\sigma(f_n)} \\ &- \left(\sum_n \frac{I_1(f_n)I_2(f_n) - I_3(f_n)(I_1(f_n)}{\sigma(f_n)} \right. \\ &\left. + \frac{I_2(f_n)) + I_3(f_n)^2}{\sigma(f_n)} \right)^2.\end{aligned}$$

As there exist only three relevant combinations of frequency placement (one sampling point at each indicator, all points at one indicator and two sampling points at one indicator, one at one of the others, see Figure 3), the above expression can be easily checked for maxima by hand. As can be seen in Figure 3, the maximum FIM-determinant is achieved when placing one of the sampling frequencies at each indicator. All other placements yield a smaller determinant. This shows that, for this example, the D-optimal design according to the developed theory is indeed in line with the vague notion of choosing sampling frequencies that "cover all important areas", and yields expected, useful designs. In contrast, looking at the A-optimality criterion

$$\text{tr}(\boldsymbol{I}(\boldsymbol{f})) = \sum_n \frac{(I_1(f_n) - I_3(f_n))^2}{\sigma(f_n)} + \sum_n \frac{(I_2(f_n) - I_3(f_n))^2}{\sigma(f_n)},$$

we can see that this criterion does not make a distinction between the cases depicted in Figure 3b and 3c. Even more, it attains its maximum for the sampling strategy depicted in figure 3d, where all sensors sample the same single spectrum. While the latter is caused by enforcing constraint (1) onto the system of equations, the first observation is valid even for an unconstrained model. Both can be seen as a result of the trace as a matrix measure not taking into account cross-dependencies of the parameters to estimate. This insight - even if of slightly heuristic nature - strengthens our justification for the statements made in III-B on D-optimal designs corresponding to uncertainty minimization under the present modeling assumptions.

V. Algorithms and Numerical Considerations

The above example shows that D-optimal designs are a promising concept for sampling frequency selection. In order to adapt the presented model to realistic scenarios, we now present algorithmic solutions for finding such D-optimal designs in more general settings.

A. Discretization and Combinatorial FIM Formulation

In order to attain solutions for given measured problem data (spectral characteristics of the form presented in II), a discretization of the frequency domain is needed. In the following, we assume that the frequency domain of interest (the range from which we are able to choose sampling frequencies) is divided into N^+ equidistant candidate frequencies $f_1^c, \ldots, f_{N^+}^c$. With this, we turn the continuous optimization problem presented in II into a combinatorial one, and thereby enable not only the use of real, measured spectral data but also combinatorial approaches on frequency selection, that are more easily able to avoid the problem of sensor clustering [14]. The discretization together with the structure of (3) enable us to reformulate the problem as a boolean optimization:

$$\begin{aligned}&\max_{\boldsymbol{\nu}} \det\left(\sum_{k=1}^{N^+} \nu_k \boldsymbol{I}_k\right), \\ &\text{s.t.} \quad \boldsymbol{\nu_k} \in \{0,1\}^{N^+} \\ &\sum_{k=1}^{N^+} \nu_k = N\end{aligned} \tag{4}$$

where $\boldsymbol{I}_k$ denotes the FIM of a single sampling frequency f_k,

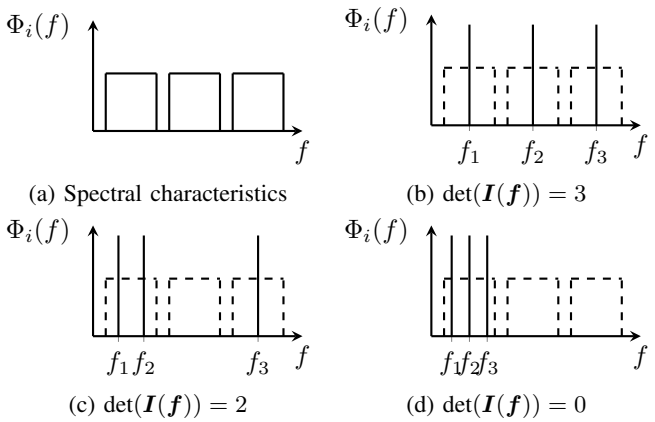
$$\boldsymbol{I}_k = \left(\frac{D_i(f_k) D_j(f_k)}{\sigma(f_k)} \right)_{i,j}.$$

To see how this discretization works on the example presented in IV, consider 5 candidate frequencies $f_1, \ldots, f_5$, where $f_2, ..., f_3$ lie each on one of the indicators, and f_1, f_5 do not cover any of the areas where one of the spectra is non-zero. Then the single-frequency FIMs defined in (V-A) become

$$\begin{aligned}\boldsymbol{I}_1 &= \boldsymbol{0} \\ \boldsymbol{I}_2 &= \text{diag}(1,0,0) \\ \boldsymbol{I}_3 &= \text{diag}(0,1,0) \\ \boldsymbol{I}_4 &= \text{diag}(0,0,1) \\ \boldsymbol{I}_5 &= \boldsymbol{0}.\end{aligned}$$

Clearly, a brute-force search for a maximum of the determinant of all possible sums of three matrices, in this case, would yield the selection $\nu_1 = \nu_5 = 0$, $\nu_2 = \nu_3 = \nu_4 = 1$. Therefore, for this example, the discretization is able to produce a result that is in line with the continuous case presented above.

$$\boldsymbol{I}(\boldsymbol{f}) = \begin{pmatrix} \sum_{n=1}^{3}(I_1(f_n) - I_3(f_n))^2 & \sum_{n=1}^{3}(I_1(f_n) - I_3(f_n))(I_2(f_n) - I_3(f_n)) \\ \sum_{n=1}^{3}(I_1(f_n) - I_3(f_n))(I_2(f_n) - I_3(f_n)) & \sum_{n=1}^{3}(I_2(f_n) - I_3(f_n))^2 \end{pmatrix}$$



(a) Spectral characteristics (b) $\det(\boldsymbol{I}(\boldsymbol{f})) = 3$ (c) $\det(\boldsymbol{I}(\boldsymbol{f})) = 2$ (d) $\det(\boldsymbol{I}(\boldsymbol{f})) = 0$

Fig. 3: A simple example of three indicator function spectra sampled using three sampling frequencies.

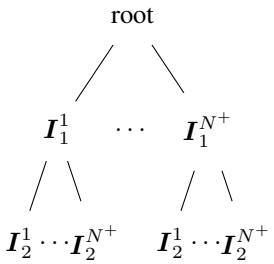


Fig. 4: Tree representation of the discretized search space of the maximum-determinant problem (4).

B. Complexity of Optimal Solutions

The formulation in (4) yields a tree-structured search space as depicted in Figure 4. Due to the nonlinearity of the determinant, for finding a maximizing combination, an exhaustive search down to the leaves of the tree requires N^{N^+} operations. Even Branch and Bound approaches as presented in [14] cannot reduce the problem to polynomial complexity. Since spectral data can be available in relatively high spectral granularity, for solving the discrete problem for larger N, approximate solutions are required to make calculations feasible, even under loose time or resource constraints.

C. Approximate Low-Complexity Approaches

1) Minkowskis Inequality: Minkowskis Inequality offers a lower bound for the problem by implying that

$$\det\left(\sum_{k=1}^{N^+}\nu_k\boldsymbol{I}_k\right) \geq \sum_{k=1}^{N^+}\det\left(\nu_k\boldsymbol{I}_k\right).$$

Maximizing this lower bound is trivial: After calculating each single frequency FIM determinant, it is as simple as picking the N candidate frequencies that yield the biggest determinants. Due to the Fisher Information Matrices being positive semi-definite and the resulting non-negative determinants, this simple algorithm is guaranteed to find the maximum of said lower bound. The resulting algorithm is in linear time. Note, however, that this bound is very loose. Even though the algorithm can find an exact solution that maximizes the lower bound, the found solution might not be sufficiently close to the real D-optimal solution in many cases. For instance, considering the example presented in Section IV, we see that all single frequency FIMs have determinant zero, and therefore the D-optimal design can not be approximated by the Minkowski-based strategy in this case. While this case might slightly degenerate in the sense that the reflection spectra have a zero value at many positions, it is reasonable to assume that, even for real spectral data, a better approximation is needed in order to attain useful results.

2) Tightening The Bound via the Entropy Power Inequality: A better approximation is motivated by the following result:

Lemma 1. *Given N symmetric PSD matrices $\boldsymbol{I}_1, \ldots, \boldsymbol{I}_N$ and some $l \leq N$ the following inequality holds:*

$$\sum_{k=1}^{N}\det(\boldsymbol{I}_k) \leq \det(\sum_{k=1}^{l-1}\boldsymbol{I}_k) + \sum_{k=l}^{N}\det(\boldsymbol{I}_k).$$

Proof:

Let $X_1, \ldots, X_N$ be independent normal random variables each distributed according to $X_k \sim \mathcal{N}(0, \boldsymbol{I}_k)$. Then

$$\begin{aligned}(2\pi\mathrm{e})\det\left(\sum_{k=1}^{N}\boldsymbol{I}_k\right)^{1/n} &= 2^{\frac{2}{n}h(X_1,\ldots,X_N)} \\ &\geq 2^{\frac{2}{n}h(X_1,\ldots,X_{N-1})} + 2^{\frac{2}{n}h(X_N)} \\ &\geq \cdots \geq 2^{\frac{2}{n}h(X_1,\ldots,X_{l-1})} + \sum_{k=l}^{N}2^{\frac{2}{n}h(X_k)} \\ &= (2\pi\mathrm{e})\det\left(\sum_{k=1}^{l-1}\boldsymbol{I}_k\right)^{1/n} + 2\pi\mathrm{e}\sum_{k=l}^{N}\det(\boldsymbol{I}_k)^{1/n},\end{aligned}$$

where the inequalities each follow from the Entropy-Power Inequality. Further $f(a+b)^{1/n} \geq f(a)^{1/n} + f(b)^{1/n}$ implies $f(a+b) \geq f(a) + f(b)$, since the logarithm is a monotone,

concave function for all $a, b > 0$, and therefore

$$\begin{aligned}\frac{1}{n}\log f(a+b) &= \log f(a+b)^{1/n} \\ &\geq \log(f(a)^{1/n} + f(b)^{1/n}) \\ &\geq \log f(a)^{1/n} + \log f(b)^{1/n} \\ &= \frac{1}{n}\log f(a) + \frac{1}{n}\log f(b))\end{aligned}$$

■

This result yields a class of bounds for all $l = 0, \ldots, N$, that are getting tighter with increasing l. Intuitively, this is a natural result in the sense that each matrix that is factored out of the sum-determinant loosens the bound to some extent. In an approximation context, this can be seen as (falsely) assuming linearity of the determinant for every matrix in $k = l, \ldots, N$, which reduces the accuracy of the approximation each time. To make the bound as tight as possible, without expanding the search space exponentially, a Greedy selection strategy can be applied. The resulting algorithm using this strategy, alongside heuristics to deal with degenerate edge cases, will be presented in the following.

3) Greedy Strategy for Better Approximate Designs: Greedy strategies are fundamental strategies for optimization based on making locally optimal choices in each iteration step. For many problems, Greedy strategies yield optimal results, e.g., if the problem to be solved has matroid structure [22]. While one of the prevalent applications of these strategies is optimization on graphs (shortest paths, maximum flows, minimum spanning trees), the graph depicted in Figure 4 and the corresponding determinant maximization problem do not satisfy the requirements for a greedy algorithm to produce optimal solutions. More specifically, the optimal solution cannot be decomposed into solutions of smaller sub-problems without approximation/loss of optimality. However, the results in V-C2 provide a way to approximate the original problem in a way that enables the use of a Greedy strategy in an efficient way: By finding

$$k_{opt,m+1} = \arg\max_{i \notin K_m} \det\left(\sum_{j \in K_m} \boldsymbol{I}_j + \boldsymbol{I}_i\right)$$

for every step m, where K_m denotes the set of indices selected up to step m, in traversing the search-space tree depicted in Figure 4, a locally optimal solution maximizing the resulting sum-determinant *so far* can be found amongst the remaining candidate frequencies at each step. Figure 5 illustrates this strategy for the first two steps of an optimization instance. The base algorithm 1 selects N sampling frequencies amongst the candidate frequencies approximating a D-optimal design in linear time.

Note, however, that this algorithm, for cases similar to the one presented in IV, i.e., for cases where (some of) the spectra are zero at some points in the sampling range, is still unable to produce meaningful results. This problem has a geometric interpretation: assume that all single-frequency FIM determinants are zero, i.e., assume no $\boldsymbol{I}_k$ to be of full rank. Then, none of the matrices are spanning the full sampling space in $\mathbb{R}^N$, and are not able to provide information about the abundance of all substances in $\mathbb{S}$ on their own. To remedy

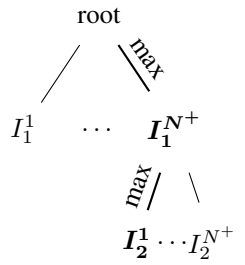


Fig. 5: Visual representation of the approximative greedy strategy motivated by the tightened Minkowski-bound on the detmax problem

Algorithm 1: Greedy-DETMAX (base algorithm)

input : Single candidate frequency FIMs $\boldsymbol{I}_k$, number N of sampling frequencies to select.
output: Indices $K = \{k_1, \ldots, k_N\}$ of the selected candidate frequencies.

$K \leftarrow \text{argmax}_k \det(\boldsymbol{I}_k)$
$i \leftarrow 1$
while $i < N$ **do**
  $k^* \leftarrow \text{argmax} \det(\sum_{j \in K} \boldsymbol{I}_j + \boldsymbol{I}_k)$
  $K \leftarrow K \cup \{k^*\}$
  $i \leftarrow i + 1$
end

this, it is integral to select indices (i.e., sampling frequencies) that move the sum-determinant closer towards having full rank. For a given set of indices already selected, to see which index to include next in order to achieve this, we can meet one of the following two scenarios:

1) At least one of the remaining FIMs increases the rank of the sum-determinant,
2) none of the remaining FIMs alone will increase the rank of the sum-determinant

Using our geometric intuition, it is clear that in case 1), we should select the next index according to

$$\arg\max_k \left(\text{rank}(\sum_{j \in K} \boldsymbol{I}_j + \boldsymbol{I}_k) - \text{rank}(\sum_{j \in K} \boldsymbol{I}_j)\right),$$

since this choice moves us closest to the sum determinant being non-zero. For case 2), another metric is needed. Making further use of the geometric interpretation given above, we can see that this case corresponds to a scenario in which no single one of the remaining sampling frequencies to select is able to capture any information about a component's spectrum so far not covered by the frequencies previously selected. However, since near the identity, from Jacobi's formula, we have

$$\det(\boldsymbol{I} + \epsilon \boldsymbol{M}) = 1 + \epsilon \text{tr}(\boldsymbol{M}) + \mathcal{O}(\epsilon^2),$$

we get a heuristic that can guide us towards desirable frequency selections even in cases in which all one-frequency

Algorithm 2: Greedy-DETMAX

input : Single candidate frequency FIMs I_k, number N of sampling frequencies to select.
output: Indices $K = \{k_1, \dots, k_N\}$ of the selected candidate frequencies.

$i \leftarrow 0$
while $i < N$ **do**
  if $\max_{k \setminus K} \det(\sum_{j \in K} \boldsymbol{I}_j + \boldsymbol{I}_k) > 0$ **then**
    $k^* \leftarrow \operatorname{argmax}_k \det(\sum_{j \in K} \boldsymbol{I}_j + \boldsymbol{I}_k)$
  else if
   $\max_k(\operatorname{rank}(\sum_{j \in K} \boldsymbol{I}_j + \boldsymbol{I}_k) - \operatorname{rank}(\sum_{j \in K} \boldsymbol{I}_j)) > 0$
   then
    $k^* \leftarrow \operatorname{argmax}_k(\operatorname{rank}(\sum_{j \in K} \boldsymbol{I}_j + \boldsymbol{I}_k) - \operatorname{rank}(\sum_{j \in K} \boldsymbol{I}_j))$
  else
    $k^* \leftarrow \operatorname{argmax}_k \operatorname{trace}(\sum_{j \in K} \boldsymbol{I}_j + \boldsymbol{I}_k)$
  $K \leftarrow K \cup \{k^*\}$
  $i \leftarrow i + 1$
end

expansions yield no increase in rank: From (V-C3), by setting $\boldsymbol{M} \equiv \boldsymbol{A} - \boldsymbol{I}$, for some small ϵ we can see that

$$\begin{aligned} \arg\max \det(\boldsymbol{A}) &\approx \arg\max(1 + \epsilon \operatorname{tr}(\boldsymbol{M})) \\ &= \arg\max(1 + \operatorname{tr}(\epsilon(\boldsymbol{A} - \boldsymbol{I}))) \\ &= \arg\max(\epsilon \operatorname{tr}(\boldsymbol{A}) - \epsilon \operatorname{tr}(\boldsymbol{I})) \\ &= \arg\max \epsilon \operatorname{tr}(\boldsymbol{A}), \end{aligned}$$

where the equalities follow from the fact that the arg max-operation is invariant to addition of and multiplication with constants, as well as the linearity of the trace. Since good designs, in general, should be close to identity in structure (see Section IV), the assumption of ϵ being small in this scenario is acceptable in the sense that this metric, used only as a final fall-back in degenerate cases with spectra that absorb 100% of the energy at some frequency, can help the algorithm move closer to reasonable designs. Putting the two problem-specific heuristics presented above together with the base algorithm yields the Greedy-DETMAX algorithm: This algorithm overcomes the weaknesses for cases like the one presented in IV, while still computing an approximate optimal design in linear time. Extensions of this algorithm can include a further look-ahead of m steps, meaning that in each iteration step i, the tuple $\{k_i, \dots, k_{i+m}\}$ is searched for an optimum by the argmax-function, i.e.,

$$k^* \leftarrow \arg \max_{k_i, \dots, k_{i+m}} \det(\sum_{j \in K} I_j + I_k)$$

for the base case. This general class simplifies to the Greedy-DETMAX algorithm for $m = 0$.

VI. Towards a Real World Use-Case: Chitin and Breeding-Place detection

To see how a frequency selection according to the outlined framework is beneficial for substance detection, this section will provide a specific setup that is close to a real-life scenario.

Mosquitoes as disease vectors are integral to the spread of common diseases like Zika and Dengue Fever in many tropical regions. These diseases pose a serious economic and humanitarian problem in those regions [23]. Successful disease control highly depends on the health authorities being able to detect mosquito breeding places. Recent research shines a light on the use of aerial drones to enable authorities to quickly and without larger personnel expenses identify and eradicate these breeding places [24], [5], [6]. One of the core problems to solve in this setting is the image classification for efficient detection of potential bodies of water carrying larval bodies.

The authors of [25] show that the presence of mosquito larvae in water causes a change in a water body's reflection spectrum. This leads us to believe that well-chosen sensor placement can enable efficient larvae detection, with algorithms being able to distinguish between shallow water bodies with larvae (positive) and without larvae (negative).

We study detection performance for a set of three substances typically found in urban environments, namely water, concrete, and sand. Additionally, as the main substance to detect for this use-case, we consider chitin layered in water: Studies suggest that chitin is one of the main substances present in larval bodies [26] and that its reflection spectrum (or the directly related absorption coefficient) shows characteristics that can enable efficient detection [27], [28]. We assume the spectrum of five layers of chitin and five layers of water each at a thickness of $75\,\mathrm{nm}$,[1] as presented in [27, Fig. 10]. We approximate the depicted curve by the biased Gaussian bell

$$\Phi_{\text{chitin}}(\lambda) = \frac{0.05}{\sqrt{0.08\pi}} \exp\left(\frac{(\lambda - 0.41)^2}{0.08}\right) + 0.05.$$

The spectral characteristics of the remaining substances are taken from [29]. All spectra are truncated to the minimum wavelength span offered by the available data of all substances and resampled to include identical grid points using SciPy's Fourier method. We assume the additive Gaussian noise model presented above with equal noise variance at every candidate frequency, i.e.,

$$\sigma(f) = \text{const.} \quad \forall f.$$

To resemble the use of a multi-spectral camera as commonly used in detection scenarios using drones, we calculate an approximate D-optimal design for $N = 5$ sampling frequencies using the proposed Greedy-DETMAX algorithm and compare the performance of abundance estimation to estimation using the 5 sampling frequencies of a MikaSense RedEdge-MX (RX02) multi-spectral sensor [30], as well as standard RGB sampling ($N = 3$). The resulting wavelengths used for sampling are provided in Table I. For each of the sampling strategies, we estimate the abundances using a constrained least squares optimization according to

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} ||\boldsymbol{\Phi}^{spec}\boldsymbol{\theta} - \boldsymbol{\Phi}^{measured}||^2$$

subject to the constraints given in (1). Here, $\boldsymbol{\Phi}^{spec}$ is a matrix containing the spectral characteristic samples at the selected

[1] Note that, since this is a purely illustrative example, the choice of this specific layering was done arbitrarily. For actual employment of this technique, we suggest the usage of a measured spectral characteristic from bodies of water containing larvae, recorded in a lab setting or a field experiment.

frequencies for the specific endmembers, and $\mathbf{\Phi}^{measured}$ contains the simulated measurement data of the mixture.

A. Abundance Estimation

We first assume an equal mixture, that is,

$$\boldsymbol{\theta}_{\text{true}} = (0.25, 0.25, 0.25, 0.25)^T.$$

Figure 6 shows the used spectral characteristics of the four substances, pre-processed as described above, and the corresponding approximated D-optimal design using the Greedy-DETMAX procedure, indicated by dashed lines marking the sampling frequencies selected by the algorithm. The depicted result underlines our interpretation of D-optimal designs in this context: For a set of spectral characteristics, the frequencies are selected to sample parts of the spectrum where the substances on average differ as much as possible. As expected with this strategy, one of the frequencies selected samples the peak of the chitin bell curve, while other frequencies emphasize the difference between, e.g., concrete and sand more strongly (f_4, f_5). Figure 7 shows the resulting average mean squared error in abundance estimation using the constrained least squares estimator in (VI) and the mixture as given in (VI-A), averaged over $10,000$ estimations.

Clearly, the approximate D-optimal sampling strategy is able to improve estimation for this case, with the largest difference in mean squared error occurring in lower-noise regions. As the noise variance becomes large, the errors for the different designs converge. This is expected, since for cases in which the noise power is significantly larger than the spectral values, estimation becomes completely random, with all information possibly produced during sampling becoming negligible due to the large variance of the sampling noise.

B. Substance Detection

To see how the sampling strategies compare for single-substance detection, we consider the singular case (see Section II), in which every "mixture" is assumed to stem from only one specific substance, i.e., the parameter vector is a one-hot encoding indicating deterministic class membership. To address the fact that in this scenario we usually are not interested in all estimated abundances, but rather want as many estimations to "detect" the correct substance (i.e., the maximum estimated abundance is the abundance of the true, underlying substance present), we slightly adapt our evaluation metric and look at the probability of correct detection:

$$P[\text{"correct detection"}] \equiv P[\hat{s} = s_i | \boldsymbol{\theta}^i_{\text{true}}],\ i = 1, \ldots, 4,$$

where

$$\hat{s} = s_{\hat{k}}, \quad \hat{k} \coloneqq \arg\max \hat{\boldsymbol{\theta}}$$

for the corresponding singular true parameter vectors

$$\begin{aligned}
\boldsymbol{\theta}^1_{\text{true}} &= (1, 0, 0, 0)^T \\
\boldsymbol{\theta}^2_{\text{true}} &= (0, 1, 0, 0)^T \\
\boldsymbol{\theta}^3_{\text{true}} &= (0, 0, 1, 0)^T \\
\boldsymbol{\theta}^4_{\text{true}} &= (0, 0, 0, 1)^T
\end{aligned}$$

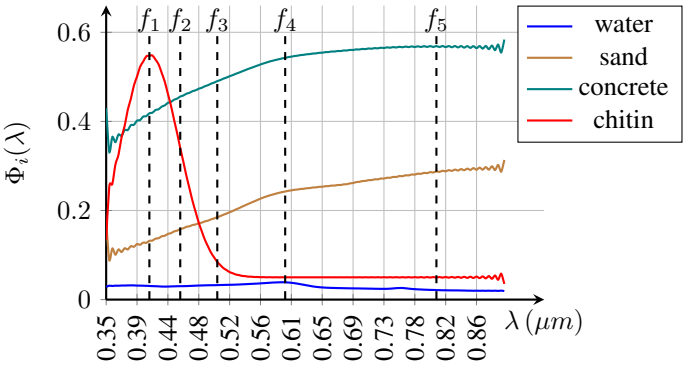


Fig. 6: Spectral characteristics of the simulated substances, including dashed markers for the sampling frequencies selected by the Greedy-DETMAX algorithm for the four substance estimation problem.

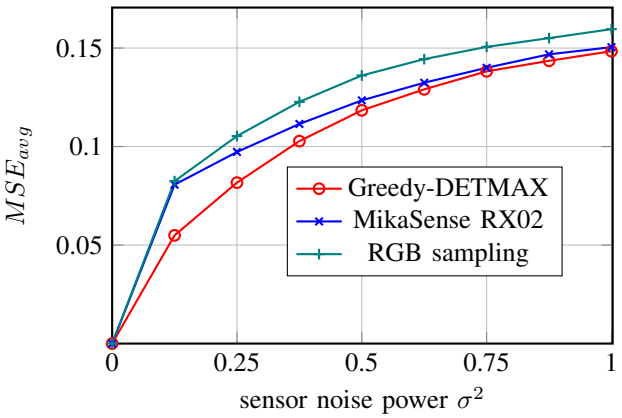
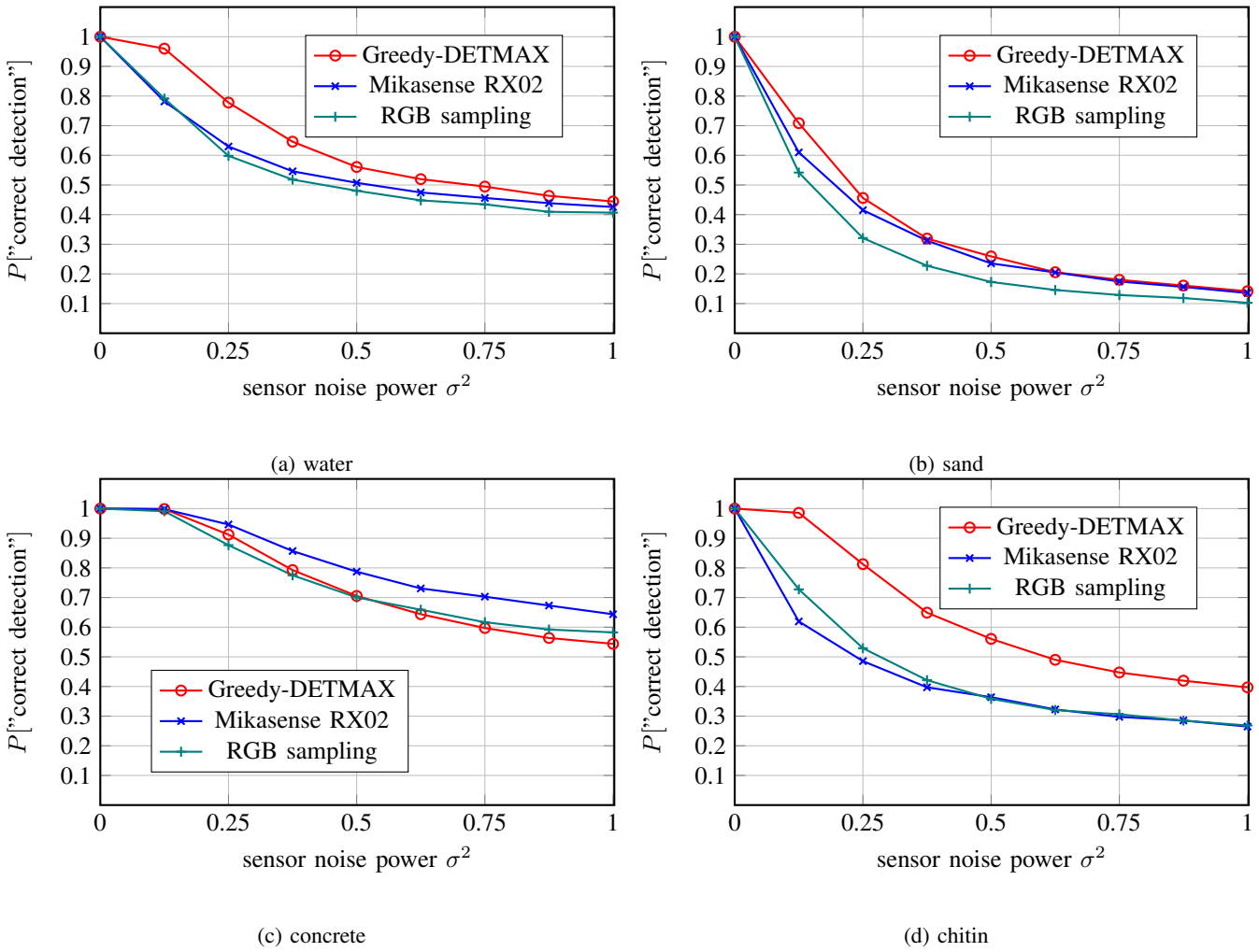


Fig. 7: Mean Squared Error averaged over 10,000 estimations of the parameter estimation in the four substance case with the abundance vector as given in (VI-A).

under increasing additive white Gaussian noise, estimated from 10,000 classifications using the constrained least squares estimator given in (VI). As can be seen in Figure 8, the probability of correct detection decreases under increasing noise power, with zero noise power enabling perfect classification in all cases, and using all sampling strategies. Further, the approximate D-optimal design can be seen to increase the probability of correct detection compared to the other sampling strategies for the classes water, sand, and chitin. Interestingly, in the case of the spectra stemming from concrete as the underlying class, the presented commercial multi-spectral sensor sampling positions outperform the D-optimal design visibly, with the D-optimal strategy coinciding almost exactly with standard RGB sampling. The largest difference in correct detection probability can be seen in Figure 8d, likely explained by the D-optimal design sampling the exact position of the rather sharp peak in the chitin spectrum, while non-optimized strategies sample at positions where the spectral information our modeled chitin reflectance provides is small.



(a) water

(b) sand

(c) concrete

(d) chitin

Fig. 8: Probability of correct detection for the four substances using the different frequency sampling designs estimated using 10,000 classifications for each noise level.

TABLE I: WAVELENGTHS CORRESPONDING TO THE SAMPLING FREQUENCIES FOR THE DIFFERENT SAMPLING STRATEGIES.

method	$\boldsymbol{\lambda_1}$	$\boldsymbol{\lambda_2}$	$\boldsymbol{\lambda_3}$	$\boldsymbol{\lambda_4}$	$\boldsymbol{\lambda_5}$
Greedy-DETMAX	0.409 µm	0.452 µm	0.502 µm	0.595 µm	0.807 µm
Mikasense RX02	0.477 µm	0.562 µm	0.667 µm	0.718 µm	0.841 µm
RBG sampling	0.464 µm	0.532 µm	0.629 µm	–	–

VII. CONCLUSION

In this paper, we derived a framework for evaluating and designing frequency sampling schemes in a multi-spectral imaging for detection context. We have provided the reader with intuition for how this framework distinguishes the quality of different schemes in line with our common sense by providing an analytically traceable example. We have extended the framework to discrete sampling frequency selection by restating the discrete optimization problem formulated in [9] for our context. We further provided novel ways of approximating optimal sampling frequency selection using lower bounds on determinant sums as well as problem-specific heuristics, that result in a Greedy-strategy greatly reducing algorithm complexity. We have shown that, for our intuitive example, this algorithm is able to find optimal solutions. Lastly, by constructing a more complex example that is closer to real substance detection, we were able to give simulated numerical evidence of our approach being beneficial in a parameter estimation problem as commonly present in detection scenarios. The presented framework can directly be used in practice to guide the choice of multi-spectral sensor configuration for a given use-case. Further, in conjunction with the presented optimization methods, the framework can guide the design of new multi-spectral sensors when considering a specific problem for which detailed spectral data is available.

Several interesting questions are open for further exploration. Firstly, the adapted Gaussian data model, while conceptually simple, likely is not the most accurate model for multispectral sensors. The given framework can easily be adapted to use other data models, e.g., the remaining two formulated in [12], and therefore constitutes a promising first step towards dealing with models closely analog to reality. In the same line of thought, the framework could be extended to factor in the sensitivity ranges of real optical sensors, by not only considering frequency points, but a small weighted frequency window around a center sampling frequency. Secondly, while a first proof-of-concept simulative work has been presented in this paper, closer inspection of reflection spectra from real measurements and the resulting designs would be of larger interest. A much more powerful statement about the increase of estimation accuracy in could then be made from, e.g., field experiments. Lastly, an analysis of the provided Greedy-strategy for sampling frequency selection and its tightness to the analytically optimal solution in a general setting would be of more general interest.

REFERENCES

[1] J. Navia, I. Mondragon, D. Patino, and J. Colorado, "Multispectral mapping in agriculture: Terrain mosaic using an autonomous quadcopter uav," in *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2016, pp. 1351–1358.

[2] W. Bouachir, K. E. Ihou, H.-E. Gueziri, N. Bouguila, and N. Bélanger, "Computer vision system for automatic counting of planting microsites using uav imagery," *IEEE Access*, vol. 7, pp. 82 491–82 500, 2019.

[3] J. Su, D. Yi, B. Su, Z. Mi, C. Liu, X. Hu, X. Xu, L. Guo, and W.-H. Chen, "Aerial visual perception in smart farming: Field study of wheat yellow rust monitoring," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2242–2249, 2021.

[4] G. Carrasco-Escobar, E. Manrique, J. Ruiz-Cabrejos, M. Saavedra, F. Alava, S. Bickersmith, C. Prussing, J. M. Vinetz, J. E. Conn, M. Moreno, and D. Gamboa, "High-accuracy detection of malaria vector larval habitats using drone-based multispectral imagery," *PLOS Neglected Tropical Diseases*, vol. 13, no. 1, pp. 1–24, 01 2019. [Online]. Available: https://doi.org/10.1371/journal.pntd.0007105

[5] D. T. Bravo, G. A. Lima, W. A. L. Alves, V. P. Colombo, L. Djogbenou, S. V. D. Pamboukian, C. C. Quaresma, and S. A. de Araujo, "Automatic detection of potential mosquito breeding sites from aerial images acquired by unmanned aerial vehicles," *Computers, Environment and Urban Systems*, vol. 90, p. 101692, 2021.

[6] A. Amarasinghe and V. B. Wijesuriya, "Drones vs dengue: A drone-based mosquito control system for preventing dengue," in *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*. IEEE, 2020, pp. 1–6.

[7] A. Huete, "11 - remote sensing for environmental monitoring," in *Environmental Monitoring and Characterization*, J. F. Artiola, I. L. Pepper, and M. L. Brusseau, Eds. Burlington: Academic Press, 2004, pp. 183–206. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780120644773500138

[8] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[9] M. Patan, *Optimal Sensor Networks Scheduling in Identification of Distributed Parameter Systems*, ser. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 2012. [Online]. Available: https://books.google.se/books?id=MHwQBwAAQBAJ

[10] T. Nagata, T. Nonomura, K. Nakai, K. Yamada, Y. Saito, and S. Ono, "Data-driven sparse sensor selection based on a-optimal design of experiment with admm," *IEEE Sensors Journal*, vol. 21, no. 13, pp. 15 248–15 257, 2021.

[11] T. C. Xygkis, G. N. Korres, and N. M. Manousakis, "Fisher information-based meter placement in distribution grids via the d-optimal experimental design," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 1452–1461, 2018.

[12] S. Ram, "Information theoretic analysis of hyperspectral imaging systems with applications to fluorescence microscopy," *Biomed. Opt. Express*, vol. 10, no. 7, pp. 3380–3403, Jul 2019. [Online]. Available: http://opg.optica.org/boe/abstract.cfm?URI=boe-10-7-3380

[13] A. Pázman, *Foundations of Optimum Experimental Design*. Springer, 1986.

[14] D. Ucinski and M. Patan, "D-optimal design of a monitoring network for parameter estimation of distributed systems," *Journal of Global Optimization*, vol. 39, pp. 291–322, 10 2007.

[15] Y. Saito, T. Nonomura, K. Yamada, K. Nakai, T. Nagata, K. Asai, Y. Sasaki, and D. Tsubakino, "Determinant-based fast greedy sensor selection algorithm," *IEEE Access*, vol. 9, pp. 68 535–68 551, 2021. [Online]. Available: https://doi.org/10.1109%2Faccess.2021.3076186

[16] A. Niculescu-Mizil and R. Caruana, "Predicting good probabilities with supervised learning," in *Proceedings of the 22nd International Conference on Machine Learning*, ser. ICML '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 625–632. [Online]. Available: https://doi.org/10.1145/1102351.1102430

[17] B. Zadrozny and C. Elkan, "Transforming classifier scores into accurate multiclass probability estimates," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 08 2002.

[18] B. R. Frieden, *Science from Fisher Information: A Unification*. Cambridge University Press, 2004.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. USA: Wiley-Interscience, 2006.

[20] F. Nielsen, "An elementary introduction to information geometry," *Entropy*, vol. 22, no. 10, p. 1100, sep 2020. [Online]. Available: https://doi.org/10.3390%2Fe22101100

[21] A. Ly, M. Marsman, J. Verhagen, R. P. Grasman, and E.-J. Wagenmakers, "A tutorial on fisher information," *Journal of Mathematical Psychology*, vol. 80, pp. 40–55, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022249617301396

[22] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms, 3rd Edition*. MIT Press, 2009. [Online]. Available: http://mitpress.mit.edu/books/introduction-algorithms

[23] M. Senanayake, S. Jayasinghe, D. Wijesundera, and M. Manamperi, "Economic cost of hospitalized non-fatal paediatric dengue at the lady ridgeway hospital for children in sri lanka," *Sri Lanka Journal of Child Health*, vol. 43, no. 4, 2014.

[24] A. Amarasinghe, C. Suduwella, L. Niroshan, C. Elvitigala, K. De Zoysa, and C. Keppetiyagama, "Suppressing dengue via a drone system," in *2017 Seventeenth International Conference on Advances in ICT for Emerging Regions (ICTer)*. IEEE, 2017, pp. 1–7.

[25] G. Carrasco-Escobar, E. Manrique, J. Ruiz-Cabrejos, M. Saavedra, F. Alava, S. Bickersmith, C. Prussing, J. M. Vinetz, J. E. Conn, M. Moreno *et al.*, "High-accuracy detection of malaria vector larval habitats using drone-based multispectral imagery," *PLoS neglected tropical diseases*, vol. 13, no. 1, p. e0007105, 2019.

[26] X. Yang, Q. Yin, Y. Xu, X. Li, Y. Sun, L. Ma, D. Zhou, and B. Shen, "Molecular and physiological characterization of the chitin synthase b gene isolated from culex pipiens pallens (diptera: Culicidae)," *Parasites & Vectors*, vol. 12, 2019.

[27] W. Vargas, D. Azofeifa, and H. Arguedas, "Refractive indices of chitin, chitosan, and uric acid with application to structural color analysis," *Optica Pura y Aplicada*, vol. 46, pp. 55–72, 03 2013.

[28] K. Madhusudhan, P. Meghana, R. Vinaya, S. Moorthy, C. Kishorekumar, H. Prakash, and V. Sivaprasad, "Extraction and characterization of chitin and chitosan from aspergillus niger, synthesis of silver-chitosan nanocomposites and evaluation of their antimicrobial potential," *Journal of advances in Biotechnology*, vol. 6, pp. 939–945, 12 2017.

[29] R. F. Kokaly, R. N. Clark, G. A. Swayze, K. E. Livo, T. M. Hoefen, N. C. Pearson, R. A. Wise, W. M. Benzel, H. A. Lowers, R. L. Driscoll, and A. J. Klein, "Usgs spectral library version 7," 2017.

[30] MikaSense, Inc., "RedEdge-MX Integration Guide," Accessed Nov. 29, 2022. [Online]. Available: https://support.micasense.com/hc/en-us/articles/360011389334-RedEdge-MX-Integration-Guide