

Photon-noise: Is a single-pixel camera better than point scanning? A  
signal-to-noise ratio analysis for Hadamard and Cosine positive  
modulation.

Supplementary Information

Camille Scotté<sup>1</sup>, Frédéric Galland<sup>1</sup>, and Hervé Rigneault<sup>1</sup>

<sup>1</sup>Aix Marseille University, CNRS, Centrale Marseille, Institut Fresnel, Marseille, France

September 30, 2022

This document contains the theoretical proofs supporting the article main text. It can nevertheless be read independently, for the reader interested in the underlying theoretical methodology and derivations.

The aim of the study is to compare the signal-to-noise ratio associated with raster-scanning and certain categories of single-pixel detection multiplexing, when the noise only arises from the photon-counting process.

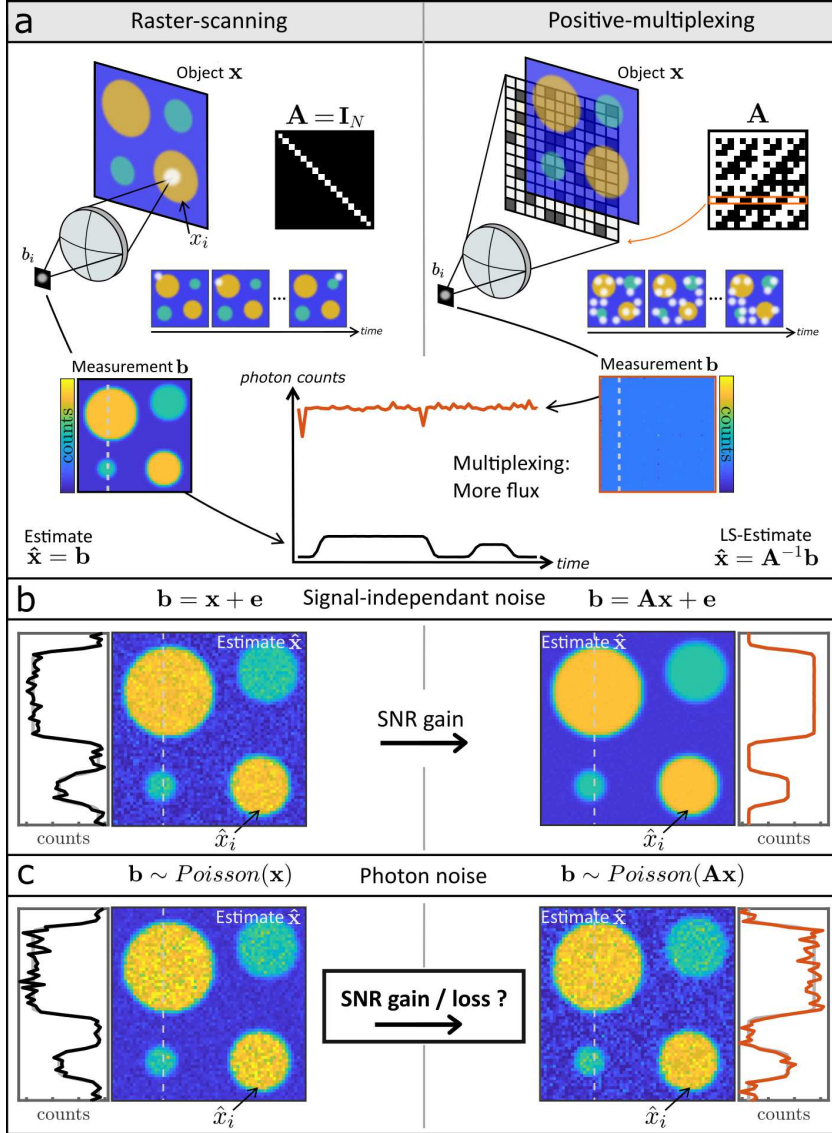


Figure 1: (a) Schematic representation of the forward model for raster-scanning and positive-multiplexing in the absence of noise. (b) Example of estimated object in the presence of additive-white Gaussian noise (signal independent noise). (c) Example of estimated object in the presence of photon noise.  $\mathbf{x}$ : intensity object,  $\hat{\mathbf{x}}$ : estimation of  $\mathbf{x}$ ,  $\mathbf{b}$ : measurement,  $\mathbf{A}$ : invertible multiplexing matrix,  $\mathbf{I}_N$ : identity matrix, LS : least-square.

In raster scanning (Fig. 1 a-left), an unknown object  $\mathbf{x}$  is scanned point by point-by-point with a focused intensity spot. In a single-pixel camera (Fig. 1 a-right), the object is sequentially illuminated by intensity patterns: the total reflected or transmitted intensity is summed into a single-pixel detector and must be demultiplexed to retrieve the object [1]. This intensity modulation multiplexing is referred to as *Positive-multiplexing*<sup>1</sup>. In both cases, because of measurement noise, the retrieved object  $\hat{\mathbf{x}}$  differs from the ground truth object  $\mathbf{x}$  by some error (Fig. 1 b-c). This error is quantified by the signal-to-noise ratio (SNR), or equivalently by the mean-square error (MSE). When the noise arises from the detector electronics (Fig. 1 b), it is well known common positive-multiplexing types dramatically improve the SNR as compared to raster-scanning [2, 3]. Yet, when the noise only arises from the photon counting process (Fig. 1 c), the answer has not been made clear, despite the development of high-performance single-pixel detectors that tends to make measurements more and more likely to be limited by photon noise only.

<sup>1</sup>because this work is also valid in the particular case of interferometric measurements where the object of interest is the field power spectrum, as will be shown in 4.3

The present document assesses the *theoretical* SNR performances of raster-scanning and positive-multiplexing. We therefore consider a general intensity object  $\mathbf{x}$  that we seek to estimate with least-square estimation (LS). We derive the theoretical SNR associated with raster-scanning and positive-multiplexing, for three measurements strategies, both for a general matrix and for Hadamard-based and Cosine-based positive-multiplexing.

The document is organised as followed: In Part 1, we define the assumptions and framework of the study. In Part 2, we derive the general expressions for the MSE associated to raster-scanning and positive-multiplexing. In Part 3, we derive the MSE for three common positive-multiplexing schemes. In Part 4, we apply these results to positive-multiplexing based on Hadamard matrices and to modulation with positive cosines. We find that for these classes of positive-multiplexing, the MSE is constant on most object pixels, and depends on the average signal contained in the object. In Part 5, we study the robustness of raster-scanning and positive-multiplexing to several perturbations. Last, in Part 6, we elaborate on some general properties of positive-multiplexing matrices that lead to a constant MSE.

In addition to multiplexing schemes, matrices, and perturbations, the SNR depends on other parameters, such as the object signal itself or the estimation strategy. The effect of the later are assessed in the article main text; which, by combining the present theoretical results with simulations and experiments, leads to implementable user guidelines. Table. 1 gives an overview of which SNR parameters are assessed in the present document and which are assessed in the article main text.

| SNR dependence                                     | Supplementary Information   | Article main text   |
|--|---|---|
| Object $\mathbf{x}$                                | General object $\mathbf{x}$   | General object $\mathbf{x}$ + Specific object classes   |
| Estimation method                                  | Least-square (LS)   | Least-square + Other estimators (LSclip, NNLS, MLE)   |
| Multiplexing scheme                                | 3 schemes: One-step, Two-step, Dual-detection   |   |
| Multiplexing matrix $\mathbf{A}$                   | General matrix $\mathbf{A}$ ; positive-Hadamard ( $\mathbf{H}_1$ -matrix and $\mathbf{S}$ -matrix), positive-Cosines ; Generalisation to certain matrix classes | positive-Hadamard ( $\mathbf{H}_1$ -matrix and $\mathbf{S}$ -matrix), positive-Cosines                            |
| Noise & measurement model                          | Photon noise, perturbations of the multiplexing matrix, perturbation of the model with various noise or nuisances   | Photon noise, perturbations of the multiplexing matrix, perturbation of the model with various noise or nuisances |
| Resolution, sampling, number of measurements, etc. | <i>Not considered in this study</i>   |   |

Table 1: Overview of some parameters influencing the SNR, and of which document assesses them.

# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Framework of the study</b>  | <b>5</b>  |
| 1.1      | General model . . . . .  | 5         |
| 1.2      | Figures-of-merit: SNR and MSE . . . . .  | 5         |
| 1.3      | Framework and assumptions . . . . .  | 6         |
| <b>2</b> | <b>Results for a general matrix <math>\mathbf{A}</math></b>                                  | <b>7</b>  |
| 2.1      | Least-square estimation (LS) . . . . .   | 7         |
| 2.2      | MSE for positive-multiplexing . . . . .  | 7         |
| 2.3      | MSE for raster-scanning . . . . .  | 8         |
| <b>3</b> | <b>Results for three positive-multiplexing schemes</b>                                       | <b>9</b>  |
| 3.1      | One-step multiplexing . . . . .  | 9         |
| 3.2      | Two-step multiplexing . . . . .  | 10        |
| 3.3      | Dual-detection . . . . .   | 10        |
| <b>4</b> | <b>Properties of positive-Hadamard-based multiplexing &amp; positive-Cosine multiplexing</b> | <b>12</b> |
| 4.1      | Main result . . . . .  | 12        |
| 4.2      | Positive-Hadamard-based multiplexing . . . . .   | 13        |
| 4.2.1    | Definition and properties of the matrices $\mathbf{S}$ and $\mathbf{H}_1$ . . . . .          | 13        |
| 4.2.2    | MSE results . . . . .  | 15        |
| 4.2.3    | Proofs of the variance results ( $\mathbf{S}$ -matrix) . . . . .                             | 16        |
| 4.2.4    | Proofs of the variance results ( $\mathbf{H}_1$ -matrix) . . . . .                           | 18        |
| 4.3      | Positive-Cosine multiplexing . . . . .   | 21        |
| 4.3.1    | Results . . . . .  | 22        |
| 4.3.2    | Proofs for Case 2 (Fig. 4.2(b)) . . . . .  | 23        |
| 4.3.3    | Proofs for Case 1 (Fig. 4.2(a)) . . . . .  | 29        |
| 4.4      | Dual detection: Justification for the balanced-strategy . . . . .                            | 29        |
| 4.5      | Numerical results . . . . .  | 30        |
| <b>5</b> | <b>Robustness to perturbations</b>   | <b>34</b> |
| 5.1      | Initial model multiplied by a constant . . . . .   | 34        |
| 5.1.1    | Implications for the three multiplexing schemes . . . . .                                    | 34        |
| 5.1.2    | Implications for comparisons at constant number of photons . . . . .                         | 35        |
| 5.2      | Initial model with additional noise or background . . . . .                                  | 36        |
| 5.2.1    | Results . . . . .  | 37        |
| 5.2.2    | Details and proofs . . . . .   | 38        |
| <b>6</b> | <b>Some facts on matrices leading to a constant variance</b>                                 | <b>43</b> |
| 6.1      | A necessary condition . . . . .  | 43        |
| 6.2      | Sufficient condition 1 . . . . .   | 45        |
| 6.3      | Sufficient condition 2 . . . . .   | 45        |

# Notations and properties

- $\mathbf{I}_N$  :  $N \times N$  identity matrix
- $\mathbf{J}_N$  :  $N \times N$  matrix of ones
- $\mathbf{1}_N$  :  $N \times 1$  vector of ones
- $\mathbf{e}_i$  :  $N \times 1$  vector of zeros with a one at index  $i$
- $\mathbf{1e}_1^T$  : matrix with ones on the first column and zeros elsewhere
- $\mathbf{e}_1\mathbf{1}^T$  : matrix with ones the first row and zeros elsewhere
- $\odot$  : Point-wise product
- $\otimes$  : Kronecker product
- $T$  : Transpose
- $*$  : Conjugate
- $H$  : Conjugate transpose
- $+$  : Pseudo-inverse
- $\hat{\mathbf{x}}$  : estimate of  $\mathbf{x}$
- $\bar{x}$  : average over all pixels of  $\mathbf{x}$
- $diag(\mathbf{A})$  : builds the vector from the diagonal elements of  $\mathbf{A}$
- $Diag(\mathbf{x})$  : builds the diagonal matrix which elements are made of the vector  $\mathbf{x}$
- $vec(\mathbf{A})$  : stacks the columns of  $\mathbf{A}$  into a vector.

In general  $\mathbf{A}$  denotes a matrix ,  $\mathbf{a}$  a vector and  $a$  a scalar (Except for the variance denoted  $\mathbf{V}$ ).

We also make extensive use of the following properties [4]:

$$diag(\mathbf{A}Diag(\mathbf{b})\mathbf{C}) = (\mathbf{A} \odot \mathbf{C}^T)\mathbf{b} \quad (\text{for real quantities}) \quad (0.0.1)$$

$$diag(\mathbf{A}Diag(\mathbf{b})\mathbf{C}) = (\mathbf{A} \odot \mathbf{C}^*)\mathbf{b} \quad (\text{for complex quantities}) \quad (0.0.2)$$

$$(\mathbf{A} + \mathbf{bc}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{bc}^T\mathbf{A}^{-1}}{1 + \mathbf{c}^T\mathbf{A}^{-1}\mathbf{b}} \quad (0.0.3)$$

$$vec(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})vec(\mathbf{X}) \quad (0.0.4)$$

# Chapter 1

## Framework of the study

To begin with, we introduce the model with its associated assumptions and figures-of-merit.

### 1.1 General model

We consider the following simple linear forward model: a real positive object  $\mathbf{x}$  is measured through a matrix  $\mathbf{A}$ , leading to an ideal noiseless measurement vector  $\mathbf{b}_0$ :

$$\mathbf{b}_0 = \mathbf{A}\mathbf{x} \quad (1.1.1)$$

where  $\mathbf{b}_0 \in \mathbb{R}_+^N$  is the vector of  $N$  noiseless measures,  $\mathbf{x} \in \mathbb{R}_+^N$  is the object of interest, and  $\mathbf{A} \in \mathbb{R}_+^{N \times N}$  is the  $N \times N$  measurement matrix, made of positive coefficients  $a_{ij} \geq 0$ . In multiplexing,  $\mathbf{A}$  is built with some mixing coefficients, while in raster-scanning, it is simply equal to the identity matrix  $\mathbf{I}_N$ . When the noise only arises from the photon-counting process, the noisy measurements read:

$$\boxed{\mathbf{b} \sim \text{Poisson}(\mathbf{A}\mathbf{x})} \quad (1.1.2)$$

where  $\mathbf{b} \in \mathbb{N}^N$  is the vector of observed photon counts. Each measured number of photons  $b_i$  is a random variable whose probability law is assumed to be a Poisson distribution of mean  $\langle b_i \rangle = \langle \delta b_i^2 \rangle = b_{0i}$ , where  $\delta b_i = b_{0i} - b_i$  is the measurement error.

### 1.2 Figures-of-merit: SNR and MSE

Since the measurements  $\mathbf{b}$  are noisy, one cannot perfectly access the ground-truth object  $\mathbf{x}$  but can only estimate it. This estimate, denoted  $\hat{\mathbf{x}}$ , is directly equal to the measurements for raster-scanning, and to their demodulation for multiplexing. In both cases, it differs from  $\mathbf{x}$  by some error  $\delta \hat{\mathbf{x}} = \hat{\mathbf{x}} - \mathbf{x}$ . The aim of this work is to assess which of raster-scanning or positive-multiplexing leads to the smallest error. This is assessed via the mean-square error (MSE) and signal-to-noise ratio (SNR). Both inform on how precise and accurate is the estimate  $\hat{\mathbf{x}}$  on each object pixel  $i$ :

$$\boxed{MSE(\hat{x}_i) = \langle (\hat{x}_i - x_i)^2 \rangle \quad ; \quad SNR(\hat{x}_i) = \frac{x_i}{\sqrt{MSE(\hat{x}_i)}}} \quad (1.2.3)$$

The potential SNR gain or loss brought by multiplexing over raster-scanning can be quantified with the following ratio:

$$G_i = \frac{SNR(\hat{x}_i)_{\text{multiplex}}}{SNR(\hat{x}_i)_{\text{raster-scan}}} = \sqrt{\frac{MSE(\hat{x}_i)_{\text{raster-scan}}}{MSE(\hat{x}_i)_{\text{multiplex}}}} \quad (1.2.4)$$

If  $G_i > 1$ , multiplexing improves the SNR on pixel  $i$ , as compared to raster-scanning, and conversely. Since the SNR and MSE are directly related, and to bypass the additional dependence on the object ground-truth, in the following we only give results in terms of MSE. The rest of the present document focuses on deriving the MSE analytical expressions associated with positive-multiplexing in different scenarios, and compare them with the MSE associated with raster-scanning.

### 1.3 Framework and assumptions

Comparing the SNR performances of multiplexing and raster-scanning could be done for a multitude of scenarios. We choose to restrict our study to the following framework:

- The quantities to estimate  $\mathbf{x}$  are real and positive
- The detected quantities  $\mathbf{b}$  are positive intensities (or photon counts)
- The multiplexing matrix  $\mathbf{A}$  is real and positive
- Although we mostly focus on cases in which multiplexing is achieved via intensity modulation (i.e. when the measurement is an incoherent sum of intensities), the study also holds for the particular case of interferometric measurements where the object of interest is the field power spectrum (e.g. Fourier-transform infrared spectroscopy), see section 4.3. These four first assumptions define what we call *Positive-multiplexing*
- The considered systems are linear systems
- The measurements are assumed to be statistically independent
- The system resolution is infinitely smaller than the finest details of the considered structures
- We work at fixed sampling: the number of measurements equals the number of probed object pixels  $N$ . The multiplexing matrix  $\mathbf{A}$  is thus square. It is also assumed to be invertible
- In order to derive analytical expressions, in this document we perform all estimations with the least-square estimator. The effect estimators with positivity constrained adapted to photon-noise are studied in the article main text.
- Unless otherwise stated, the results are valid for any object dimensionality (1-D, 2-D, ...), as long as the variables can be rearranged in the form of equation (1.1.1).

Last, except in the dedicated section 5.1.2, the comparison between raster-scanning and positive-multiplexing is *not* performed at constant photon number. Rather, we compare raster scanning and positive-multiplexing for fixed exposure time and irradiance. On the example of Fig.1, this means each sample pixel is illuminated with the same light power: if in raster-scanning, each pixel of the object is illuminated with 1 mW during 1 ms, then in positive-multiplexing, each pixel of the object will also see 1 mW of incident light during 1 ms. This results in a consequently higher measured number of photons for positive-multiplexing (Fig. 1 a). As we will see, even in this advantageous case, positive-multiplexing does not always lead to a better SNR than raster-scanning.

## Chapter 2

# Results for a general matrix $\mathbf{A}$

In this section we derive the theoretical expressions for the MSE associated with raster-scanning and positive-multiplexing, from which the SNR expressions can be deduced using equation (1.2.3). For positive-multiplexing, MSE depends on the estimation method employed to demodulate the measurements. In this document, we use the least-square estimation (LS).

### 2.1 Least-square estimation (LS)

The LS estimator minimizes the squared  $l_2$  norm between the noisy and noiseless measurements:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{b} - \mathbf{b}_0\|^2 \quad (2.1.1)$$

If  $\mathbf{A}$  is invertible, the LS solution reads:

$$\hat{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{b} \quad (2.1.2)$$

This estimation is unbiased:

$$\langle \hat{\mathbf{x}} \rangle = \mathbf{A}^{-1}\langle \mathbf{b} \rangle = \mathbf{A}^{-1}\mathbf{b}_0 = \mathbf{x} \quad (2.1.3)$$

This estimator is optimal in the sense of the maximum-likelihood for additive-white Gaussian noise, but may not be optimal under photon-noise. In this document, we use the LS estimator to access an unbiased estimation and to derive the MSE.

analytical expressions.

Since the LS estimator is unbiased, the following relation holds:

$$\mathbf{MSE}(\hat{\mathbf{x}}) = \mathbf{V}(\hat{\mathbf{x}}) = \operatorname{diag}(\mathbf{\Gamma}) \quad (2.1.4)$$

The MSE is equal to the estimation variance  $\mathbf{V}(\hat{\mathbf{x}})$ , which is itself equal to the diagonal of the covariance matrix  $\mathbf{\Gamma} = \langle \delta\hat{\mathbf{x}}\delta\hat{\mathbf{x}}^T \rangle$ .

### 2.2 MSE for positive-multiplexing

We remind that the measurements read:

$$\mathbf{b} \sim \operatorname{Poisson}(\mathbf{A}\mathbf{x}) \quad (2.2.5)$$

The estimation error arising when estimating  $\hat{\mathbf{x}}$  with LS via equation (2.1.2) is:

$$\delta\hat{\mathbf{x}} = \hat{\mathbf{x}} - \mathbf{x} = \mathbf{A}^{-1}(\mathbf{b} - \mathbf{b}_0) = \mathbf{A}^{-1}\delta\mathbf{b} \quad (2.2.6)$$

Inserting this in equation (2.1.4) yields:

$$\mathbf{\Gamma} = \langle \delta\hat{\mathbf{x}}\delta\hat{\mathbf{x}}^T \rangle = \mathbf{A}^{-1}\langle \delta\mathbf{b}\delta\mathbf{b}^T \rangle\mathbf{A}^{-T} \quad (2.2.7)$$



Since the measurements are assumed to be statistically independent and to follow a Poisson distribution,  $\langle \delta \mathbf{b} \delta \mathbf{b}^T \rangle$  is a diagonal matrix with elements  $\langle \delta b_i^2 \rangle = \langle b_i \rangle = b_{0i}$  (Poisson distribution properties). Thus:

$$\mathbf{\Gamma} = \mathbf{A}^{-1} \text{Diag}(\mathbf{b}_0) \mathbf{A}^{-T} = \mathbf{A}^{-1} \text{Diag}(\mathbf{A} \mathbf{x}) \mathbf{A}^{-T} \quad (2.2.8)$$

The MSE and estimation variance are finally obtained by selecting the diagonal of  $\mathbf{\Gamma}$  (using equations (2.1.4) and (0.0.1)). Hence, the MSE associated with positive-multiplexing reads:

$$\boxed{\mathbf{MSE}(\hat{\mathbf{x}}) = \mathbf{V}(\hat{\mathbf{x}}) = (\mathbf{A}^{-1} \odot \mathbf{A}^{-1}) \mathbf{A} \mathbf{x}} \quad (2.2.9)$$

### 2.3 MSE for raster-scanning

In raster-scanning (Fig. 1 a), the object  $\mathbf{x}$  is probed point-by-point: each measurement  $b_i$  relates to one object pixel intensity  $x_i$ . Although in practice there is no multiplexing matrix, the latter can be viewed as the identity matrix  $\mathbf{A} = \mathbf{I}_N$ . The measurements read:

$$\mathbf{b} \sim \text{Poisson}(\mathbf{x}) \quad (2.3.10)$$

The object estimate is directly equal to the noisy measurement:

$$\hat{\mathbf{x}} = \mathbf{b} \quad (2.3.11)$$

and the error is  $\delta \hat{\mathbf{x}} = \delta \mathbf{b}$ . Applying (2.2.9) with  $\mathbf{A} = \mathbf{I}_N$  gives:

$$\boxed{\mathbf{MSE}(\hat{\mathbf{x}}) = \mathbf{V}(\hat{\mathbf{x}}) = \mathbf{x}} \quad (2.3.12)$$

In other words, in raster-scanning, the MSE (or variance) is equal to the object pixel intensity on every pixel  $i$ . Therefore, from equation (1.2.3), in raster-scanning the SNR reads  $\text{SNR}(\hat{\mathbf{x}}) = \sqrt{\mathbf{x}}$ .

In the following, we only use LS-estimation, therefore we mostly speak about variance since:

$$\mathbf{MSE}(\hat{\mathbf{x}}) = \mathbf{V}(\hat{\mathbf{x}}) \quad (2.3.13)$$

which is related to the SNR via equation (1.2.3).

## Chapter 3

# Results for three positive-multiplexing schemes

In practise, positive-multiplexing can be implemented in many different ways, which may impact the final SNR. In this section, we detail three possible positive-multiplexing schemes. We associate each scheme with its equivalent multiplexing matrix and theoretical variance. The figure here-in-below encompasses the main results of this section. Explanations and details are given in the text which follows.

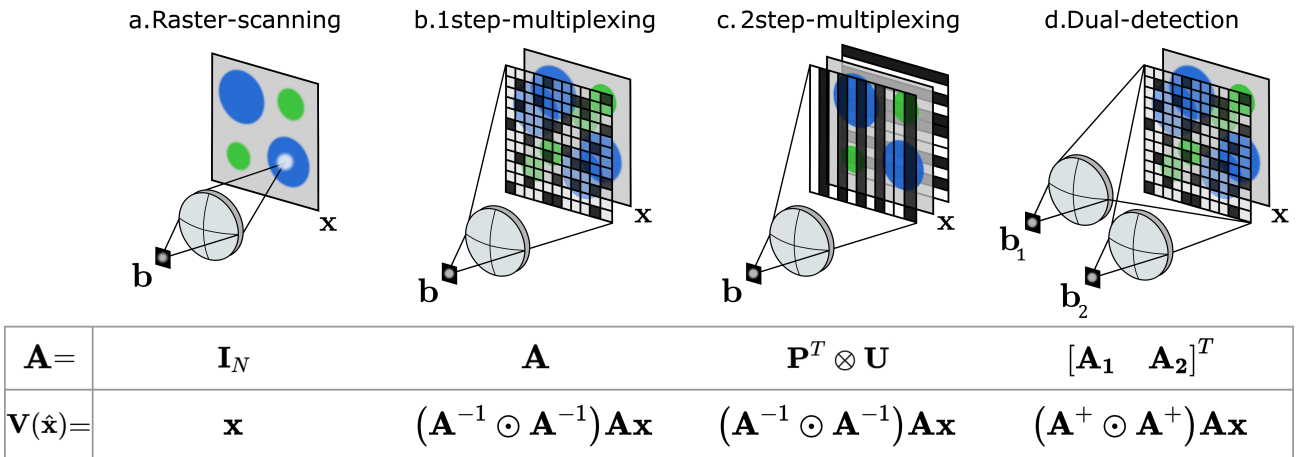


Figure 3.1: Schematic of 2-D multiplexing for raster-scanning and for the considered multiplexing modalities. To each scheme corresponds an equivalent matrix  $\mathbf{A}$  and estimation variance  $\mathbf{V}(\hat{\mathbf{x}})$  ( $= \mathbf{MSE}(\hat{\mathbf{x}})$ ). All details are given in the text. (b) In one-step multiplexing, the patterns derive from the rows of  $\mathbf{A}$ . (c) Two-step multiplexing involves two independent sets of patterns: one set derives from the rows of a matrix  $\mathbf{U}$  and the other from the rows of a matrix  $\mathbf{P}$  (see Fig. 3.2 c). (d) Dual-detection involves two complementary measurements. The equivalent matrix is a bloc matrix (equation (3.3.8)). In this scheme, it is sometimes possible to subtract the two measurements but this leads to a different model and theoretical variance (section 3.3).

### 3.1 One-step multiplexing

In one-step multiplexing (Fig. 3.1 b), the object is probed with patterns corresponding to its dimensionality. As detailed in Fig. 3.2 (a-b), each measurement  $b_i$  is the sum of the point-wise product between the object and a pattern that derives from the  $i^{th}$  row of the multiplexing matrix  $\mathbf{A}$ . The variance (or MSE) is the same as in equation (2.2.9):

$$\mathbf{V}(\hat{\mathbf{x}}) = (\mathbf{A}^{-1} \odot \mathbf{A}^{-1}) \mathbf{A} \mathbf{x} \quad (3.1.1)$$

### 3.2 Two-step multiplexing

If the object has more than one dimension, it is possible to perform multiplexing in one single step (as above), or in several steps, if its dimensions are separable. In the case of a 2-D object, this would involve two independent 1-D multiplexing stages that probe uncorrelated dimensions of the object, such as the vertical and horizontal dimensions of a 2-D spatial object (Fig. 3.1 c), or the spatial and spectral dimensions of a spatio-spectral object. Then, instead of multiplexing with a matrix  $\mathbf{A}$  of dimension  $N \times N$ , multiplexing is performed independently by two matrices  $\mathbf{P}$  and  $\mathbf{U}$ , each of size  $\sqrt{N} \times \sqrt{N}$ , as detailed in Fig. 3.2(c). Writing  $\mathbf{x} = \text{vec}(\mathbf{X})$ , and using (0.0.4) allows writing the measure as:

$$\mathbf{b}_0 = \text{vec}(\mathbf{UXP}) = (\mathbf{P}^T \otimes \mathbf{U})\mathbf{x} = \mathbf{A}\mathbf{x} \quad (3.2.2)$$

and

$$\mathbf{b} \sim \text{Poisson}((\mathbf{P}^T \otimes \mathbf{U})\mathbf{x}) = \text{Poisson}(\mathbf{A}\mathbf{x}) \quad (3.2.3)$$

Hence, this is equivalent to multiplexing with an equivalent matrix:

$$\mathbf{A} = \mathbf{P}^T \otimes \mathbf{U} \quad (3.2.4)$$

leading to an equivalent variance derived from equation (2.2.9):

$$\begin{aligned} \mathbf{V}(\hat{\mathbf{x}}) &= ((\mathbf{P}^T \otimes \mathbf{U})^{-1} \odot (\mathbf{P}^T \otimes \mathbf{U})^{-1}) (\mathbf{P}^T \otimes \mathbf{U})\mathbf{x} \\ &= ((\mathbf{P}^{-T} \odot \mathbf{P}^{-T})\mathbf{P}^T \otimes (\mathbf{U}^{-1} \odot \mathbf{U}^{-1})\mathbf{U})\mathbf{x} \end{aligned} \quad (3.2.5)$$

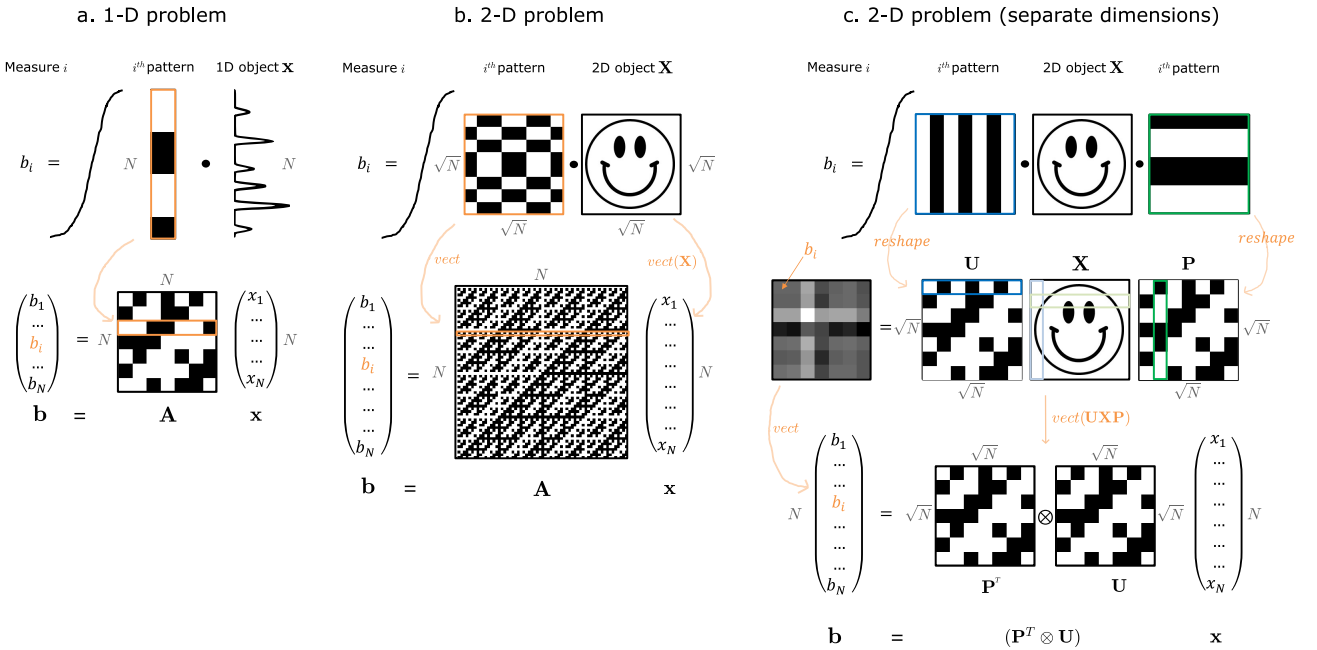


Figure 3.2: (a) One-step multiplexing in 1-D: the  $i^{\text{th}}$  row of  $\mathbf{A}$  is directly the  $i^{\text{th}}$  pattern (b) One-step multiplexing in 2-D: the  $i^{\text{th}}$  row of  $\mathbf{A}$  is reshaped to obtain a 2-D pattern. (c) Two-step multiplexing: if the object dimensions are separable and probed independently, multiplexing in 2-D can be performed in 2-steps, what we call two-step multiplexing.

### 3.3 Dual-detection

A dual detection scheme (Fig. 3.1 d) involves two complementary measurements  $\mathbf{b}_1$  and  $\mathbf{b}_2$ . For example, when multiplexing 0 and 1 entries, the signal corresponding to zeros not collected by the first detector is collected by the second detector. This is equivalent to associating a multiplexing matrix  $\mathbf{A}_1$  to the measurements  $\mathbf{b}_1$  and multiplexing matrix  $\mathbf{A}_2$  to the measurements  $\mathbf{b}_2$ , such that:

$$\mathbf{A}_1 + \mathbf{A}_2 = \mathbf{J}_N \quad (3.3.6)$$

The measurements  $\mathbf{b}_1$  and  $\mathbf{b}_2$  associated with matrices  $\mathbf{A}_1$  and  $\mathbf{A}_2$  respectively read:

$$\begin{cases} \mathbf{b}_1 \sim \text{Poisson}(\mathbf{A}_1 \mathbf{x}) \\ \mathbf{b}_2 \sim \text{Poisson}(\mathbf{A}_2 \mathbf{x}) \end{cases} \quad (3.3.7)$$

The measurement model is obtained by combining the two measurements into the same vector:

$$\mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} \sim \text{Poisson} \left( \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{x} \right) = \text{Poisson}(\mathbf{A} \mathbf{x}) \quad (3.3.8)$$

where  $\mathbf{A}$  is a block matrix of size  $2N \times N$ . It can be seen as an equivalent multiplexing matrix:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{bmatrix}^T \quad (3.3.9)$$

Then, if  $\mathbf{A}^T \mathbf{A}$  is not singular,  $\hat{\mathbf{x}}$  can be estimated via  $\hat{\mathbf{x}} = \mathbf{A}^+ \mathbf{b}$ , where  $\mathbf{A}^+$  is the pseudo-inverse:  $\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ . Replacing in equation (2.2.9) leads to the associated estimation variance:

$$\boxed{\mathbf{V}(\hat{\mathbf{x}}) = (\mathbf{A}^+ \odot \mathbf{A}^+) \mathbf{A} \mathbf{x}} \quad (3.3.10)$$

### Alternative approach: the balanced approach

In the literature (eg. [5–8]), it is often found that the reconstituted measure  $\mathbf{b}$  is rather expressed as:

$$\mathbf{b} = \mathbf{b}_1 - \mathbf{b}_2 \sim \text{Poisson}(\mathbf{A}_1 \mathbf{x}) - \text{Poisson}(\mathbf{A}_2 \mathbf{x}) \neq \text{Poisson}(\mathbf{A} \mathbf{x}) \quad (3.3.11)$$

Therefore, it does not follow the prerequisite measurement model of equation (2.2.5). Still, it is possible to define an equivalent multiplexing matrix  $\mathbf{A}$ :

$$\mathbf{A} = \mathbf{A}_1 - \mathbf{A}_2 \quad (3.3.12)$$

and to perform the estimation through:  $\hat{\mathbf{x}} = \mathbf{A}^{-1} \mathbf{b}$ , if  $\mathbf{A}$  is invertible. This is the strategy used in the literature cited above. Such an estimation strategy may seem surprising since it may lead to a significant information loss. Yet, *in some particular cases*, this strategy is justified because it is computationally efficient and leads to approximately the same variance than when using the model of equation (3.3.8) (see section 4.4 for more details). To derive the variance associated with this balanced approach, one cannot directly apply equation (2.2.9). Rather, one needs to go back to the definition of the covariance matrix, and finds:

$$\boxed{\mathbf{V}(\hat{\mathbf{x}}) = N \bar{x} (\mathbf{A}^{-1} \odot \mathbf{A}^{-1}) \mathbf{1}_N} \quad (3.3.13)$$

Proof: Covariance matrix:

$$\begin{aligned} \mathbf{\Gamma} &\equiv \langle \delta \hat{\mathbf{x}} \delta \hat{\mathbf{x}}^T \rangle \\ &= \mathbf{A}^{-1} \langle \delta \mathbf{b} \delta \mathbf{b}^T \rangle \mathbf{A}^{-T} \\ &= \mathbf{A}^{-1} \langle (\mathbf{b} - \mathbf{b}_0) (\mathbf{b} - \mathbf{b}_0)^T \rangle \mathbf{A}^{-T} \\ &= \mathbf{A}^{-1} \langle (\delta \mathbf{b}_1 - \delta \mathbf{b}_2) (\delta \mathbf{b}_1 - \delta \mathbf{b}_2)^T \rangle \mathbf{A}^{-T} \end{aligned} \quad (3.3.14)$$

Since the measurements are statistically independent,  $\langle \delta \mathbf{b}_1 \delta \mathbf{b}_2^T \rangle = \langle \delta \mathbf{b}_2 \delta \mathbf{b}_1^T \rangle = 0$ , and:

$$\begin{aligned} \mathbf{\Gamma} &= \mathbf{A}^{-1} (\langle \delta \mathbf{b}_1 \delta \mathbf{b}_1^T \rangle + \langle \delta \mathbf{b}_2 \delta \mathbf{b}_2^T \rangle) \mathbf{A}^{-T} \\ &= \mathbf{A}^{-1} \text{Diag}(\mathbf{b}_{1,0} + \mathbf{b}_{2,0}) \mathbf{A}^{-T} \\ &= \mathbf{A}^{-1} \text{Diag}((\mathbf{A}_1 + \mathbf{A}_2) \mathbf{x}) \mathbf{A}^{-T} \\ &= \mathbf{A}^{-1} \text{Diag}(\mathbf{J}_N \mathbf{x}) \mathbf{A}^{-T} = N \bar{x} (\mathbf{A}^T \mathbf{A})^{-1} \end{aligned} \quad (3.3.15)$$

Using equations (2.1.4) and (0.0.1) leads to the variance expression of equation (3.3.13).

## Chapter 4

# Properties of positive-Hadamard-based multiplexing & positive-Cosine multiplexing

So far, we have derived the expressions of the MSE for raster-scanning and three positive-multiplexing schemes, for a general multiplexing matrix  $\mathbf{A}$  verifying several assumptions (section 1.3). In the present section, we derive the MSE for two widely implemented positive-multiplexing types, based on (i) positive-Hadamard multiplexing and (ii) positive-Cosine multiplexing.

On the one hand, we consider positive-multiplexing based on Hadamard matrices, i.e. on binary matrices derived from the Hadamard matrix  $\mathbf{H}$ . We consider two specific cases: the  $\mathbf{S}$ -matrix (section 4.2.3) and the positive-Hadamard matrix  $\mathbf{H}_1$  (section 4.2.4). On the other hand, we consider positive-Cosine multiplexing, i.e. modulation with cosine waveforms (section 4.3).

| Multiplexing class | positive-Hadamard-based                                    |   | positive-Cosine              |
|--------------------|--|---|------------------------------|
| Associated matrix  | $\mathbf{S}$ -matrix                                       | $\mathbf{H}_1$ -matrix                            | $\mathbf{C}_1$ -matrix       |
| Short description  | modified Hadamard matrix with binary coefficients (0 or 1) | Hadamard matrix with binary coefficients (0 or 1) | matrix with cosine waveforms |

Table 4.1: Summary of the multiplexing classes considered in this section.

This section is organised as followed: (i) main result; (ii) positive-Hadamard-based multiplexing (results, matrices and proofs); (iii) positive-Cosine multiplexing (results, details, matrices and proofs); (iv) note on dual-detection; (v) few numerical simulations to illustrate the results.

### 4.1 Main result

We prove that for both positive-Hadamard-based multiplexing and positive-Cosine multiplexing and for three positive-multiplexing scheme, the estimation variance obtained with least-square estimation is constant over the estimated object  $\hat{\mathbf{x}}$  on most pixels  $i$ . We show that it is proportional to the average signal contained in the object  $\bar{x}$ , on *most* object pixels  $i$ :

$$\boxed{MSE(\hat{x}_i) = V(\hat{x}_i) \approx k\bar{x}} \text{ if } N \gg 1 \quad (4.1.1)$$

where  $N$  is the number of pixels and  $k$  is a positive constant. Since the variance associated with raster-scanning equals the object itself (equation (2.3.12)), positive-multiplexing improves the SNR

over raster-scanning by a factor (equation (1.2.4)) equal to:

$$G_i = \sqrt{\frac{x_i}{k\bar{x}}} \quad (4.1.2)$$

This means that, under the assumptions considered in this work, positive-multiplexing brings an improvement over raster-scanning only on object pixels  $i$  that verify:

$$\boxed{x_i \geq k\bar{x}} \quad (4.1.3)$$

The direct consequence is that the considered multiplexing strategies *do not* systematically bring an improvement over raster-scanning: it improves the SNR only on object pixels  $i$  with an intensity greater than the threshold value  $k\bar{x}$ . This result is of primary importance when choosing an optical design or a measurement strategy. The constant  $k$  depends on the multiplexing matrix, on multiplexing modality (one-step multiplexing, two-step multiplexing, dual detection), and on some specificities detailed below.

Note: The matrices presented in this section are defined so that they are directly implementable on physical systems. This implies that they do not necessarily have the same matrix norm <sup>1</sup>. To adjust the constant  $k$  for matrices norms or for multiplicative constants, refer to section 5.1.

## 4.2 Positive-Hadamard-based multiplexing

First, we focus on positive-multiplexing based on Hadamard matrices, based on binary matrices derived from the Hadamard matrix  $\mathbf{H}$ . In practise, this Hadamard-based positive multiplexing is often implemented by modulating the light intensity with encoding patterns (e.g. with a light modulator device) [3, 9–13], such as in Fig. 1 a. We consider two multiplexing matrices that are both based on Hadamard-multiplexing: the  $\mathbf{S}$ -matrix (section 4.2.3) and the positive-Hadamard matrix  $\mathbf{H}_1$  (section 4.2.4).

### 4.2.1 Definition and properties of the matrices $\mathbf{S}$ and $\mathbf{H}_1$

Positive-multiplexing based on the Hadamard matrices makes use of modified Hadamard matrices such that the matrix entries are 0 and +1 rather than  $-1$  or  $+1$ . In this work, we consider :

- The  $\mathbf{H}_1$ -matrix (denoted as "positive-Hadamard matrix") is a binary Hadamard matrix where the  $-1$  elements of  $\mathbf{H}$  are replaced with 0s.
- The  $\mathbf{S}$ -matrix is a binary matrix defined by (4.2.5). A  $\mathbf{S}$ -matrix of size  $N \times N$  can for instance be obtained by removing the first row and column of a  $(N+1) \times (N+1)$  Hadamard matrix, and changing its +1s to 0s and  $-1$ s to +1s.

### The Hadamard matrix

A Hadamard matrix  $\mathbf{H}$  is a real square matrix with entries  $-1$  and  $1$  whose rows are pairwise orthogonal [9, 14, 15]. The orthogonality condition means that the dot product of any two distinct rows is zero; and it implies that, when comparing two rows, the number of matchings ( $+1$ ) is equal to the number of mismatches ( $-1$ ). A Hadamard matrix  $\mathbf{H}$  of size  $N \times N$  is such that [9, 15]:

$$\mathbf{H}\mathbf{H}^T = \mathbf{H}^T\mathbf{H} = N\mathbf{I}_N \quad (4.2.4)$$

Hadamard matrices do not exist for any  $N$ , but it is conjectured that there is at least one Hadamard matrix of order  $N = 4p$  for every positive integer  $p$  [16]. There are several subcategories of Hadamard matrices. In this work, we consider the widely used Sylvester type [14]. For Sylvester Hadamard

---

<sup>1</sup>Largest singular value

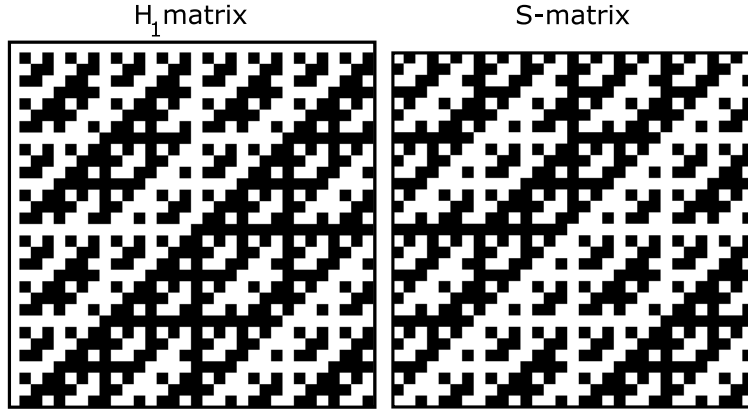


Figure 4.1:  $\mathbf{H}_1$ : Positive-Hadamard matrix ( $32 \times 32$ ), and corresponding  $\mathbf{S}$ -matrix ( $31 \times 31$ ), with  $(0,1) \equiv (\text{black}, \text{white})$

matrices,  $N = 2^p$  and  $\mathbf{H}$  is a symmetric matrix ( $\mathbf{H}^T = \mathbf{H}$ ) in which each row and column contains the same number of  $+1$  and  $-1$  elements, except for the first row and column where all elements are  $+1$ s.

Note: under additive white Gaussian noise (AWGN), Hadamard matrices provide an "optimal encoding design": among matrices which elements can be  $-1, 0$  or  $1$ , Hadamard matrices minimize the MSE [9, 15].

### The $\mathbf{S}$ -matrix

A  $\mathbf{S}$ -matrix of size  $N \times N$  is a binary matrix defined by [9]:

$$\begin{cases} \mathbf{S}^T \mathbf{S} = \mathbf{S} \mathbf{S}^T = \frac{N+1}{4} (\mathbf{I} + \mathbf{J}_N) \\ \mathbf{J}_N \mathbf{S} = \mathbf{S} \mathbf{J}_N = \frac{N+1}{2} \mathbf{J}_N \end{cases} \quad (4.2.5)$$

This means that the sum of each column (or row) of a  $\mathbf{S}$ -matrix is equal to  $\frac{N+1}{2}$ . In this work, the  $\mathbf{S}$ -matrix of size  $N \times N$  is obtained by removing the first row and column of a  $(N+1) \times (N+1)$  Sylvester Hadamard matrix, and changing its  $+1$ s to  $0$ s and  $-1$ s to  $+1$ s. In these conditions,  $\mathbf{S}$  is of odd dimension, symmetric ( $\mathbf{S} = \mathbf{S}^T$ ), and invertible with:

$$\mathbf{S}^{-1} = \frac{2}{N+1} (2\mathbf{S}^T - \mathbf{J}_N) \quad (4.2.6)$$

and  $\mathbf{S}^{-1}$  verifies:

$$\mathbf{S}^{-1} \odot \mathbf{S}^{-1} = \frac{4}{(N+1)^2} \mathbf{J}_N \quad (4.2.7)$$

Note: under additive white Gaussian noise,  $\mathbf{S}$ -matrices minimize the MSE among matrices with entries  $0$  and  $1$  [9, 17, 18].

### The $\mathbf{H}_1$ -matrix

The  $\mathbf{H}_1$ -matrix can be expressed as a function of  $\mathbf{H}$  and  $\mathbf{J}_N$  (constant matrix made of  $+1$  elements):

$$\mathbf{H}_1 = \frac{1}{2} (\mathbf{J}_N + \mathbf{H}) \quad (4.2.8)$$

$\mathbf{H}_1$  is invertible, and its invert can be expressed analytically via equation (0.0.3):

$$\mathbf{H}_1^{-1} = \frac{2}{N} (\mathbf{H}^T - \frac{N}{2} \mathbf{e}_1 \mathbf{e}_1^T) \quad (4.2.9)$$

We denote  $\mathbf{H}_2$  the complementary matrix of  $\mathbf{H}_1$  ('negative' Hadamard matrix where the +1 elements of  $\mathbf{H}$  are replaced with 0s and -1 element with +1s):

$$\mathbf{H}_2 = \frac{1}{2}(\mathbf{J}_N - \mathbf{H}) \quad (4.2.10)$$

$\mathbf{H}_2$  is not invertible. Note that, in addition to (4.2.4), a Sylvester Hadamard matrix has, among others, the following properties:

$$\mathbf{H} \odot \mathbf{H} = \mathbf{J}_N \quad (4.2.11)$$

$$\mathbf{J}_N \mathbf{H} = N \mathbf{1e}_1^T \quad (4.2.12)$$

$$\mathbf{H} \mathbf{J}_N = N \mathbf{e}_1 \mathbf{1}^T \quad (4.2.13)$$

$$\mathbf{1e}_1^T \mathbf{H} = \mathbf{J}_N \quad (4.2.14)$$

$$\mathbf{H} \mathbf{1e}_1^T = N \mathbf{e}_1 \mathbf{e}_1^T \quad (4.2.15)$$

#### 4.2.2 MSE results

For both matrices, the results of the associated estimation variances (or MSE), for LS-estimation, are given in the Tables below. The proofs are given in sections 4.2.3 and 4.2.4.

|                  | One-step multiplexing        | Two-step multiplexing           | Dual-detection                  |                              |
|------------------|------------------------------|---------------------------------|---------------------------------|------------------------------|
| <b>A</b> -matrix | $\mathbf{S}$                 | $\mathbf{S} \otimes \mathbf{S}$ | $[\mathbf{S} \ \mathbf{S}_2]^T$ | $\mathbf{S} - \mathbf{S}_2$  |
| $V(\hat{x}_i)$   | $2\bar{x} \quad (\forall i)$ | $4\bar{x} \quad (\forall i)$    | $\bar{x} \quad (\forall i)$     | $2\bar{x} \quad (\forall i)$ |
| $k$              | 2                            | 4                               | 1                               | 2                            |

Table 4.2: Estimation variances associated to the  $\mathbf{S}$ -matrix, for three multiplexing-schemes. The variance expressions on pixel  $i$  are valid for large number of pixels  $N \gg 1$  (exact expressions in the proofs). **A**-matrix: Global multiplexing matrix.  $\mathbf{S}_2$ : complementary matrix such that  $\mathbf{S} + \mathbf{S}_2 = \mathbf{J}_N$ . Last column: balanced-detection strategy (section 3.3.11).  $\bar{x}$ : object intensity average.  $N$ : object size.

|                  | One-step multiplexing   | Two-step multiplexing   | Dual-detection                    |                               |
|------------------|---|---|-----------------------------------|-------------------------------|
| <b>A</b> -matrix | $\mathbf{H}_1$  | $\mathbf{H}_1 \otimes \mathbf{H}_1$                             | $[\mathbf{H}_1 \ \mathbf{H}_2]^T$ | $\mathbf{H}_1 - \mathbf{H}_2$ |
| $V(\hat{x}_i)$   | $2\bar{x} \quad (\forall i \neq 1)$<br>$(N - 2)\bar{x} \quad (i=1)$ | $4\bar{x} \quad (\forall i \neq n_1)$<br>$c', d' \quad (i=n_1)$ | $\bar{x} \quad (\forall i)$       |                               |
| $k$              | 2   | 4   | 1                                 |                               |

Table 4.3: Estimation variances associated to the positive-Hadamard matrix  $\mathbf{H}_1$ , for three multiplexing-schemes. The variance expressions on pixel  $i$  are valid for large number of pixels  $N \gg 1$  (exact expressions in the proofs). The results in black are valid on most pixels, while the results in gray are valid only on specific pixels ( $n_1$ : pixels on the first row and column of the image,  $c', d'$ : values given in equation (4.2.35)). **A**-matrix: Global multiplexing matrix.  $\mathbf{H}_2$ : complementary matrix such that  $\mathbf{H}_1 + \mathbf{H}_2 = \mathbf{J}_N$ . Last column: balanced-detection strategy (section 3.3.11).  $\bar{x}$ : object intensity average.  $N$ : object size.



Therefore, for a given one-step multiplexing scheme, positive-Hadamard multiplexing with the  $\mathbf{S}$ -matrix or the  $\mathbf{H}_1$ -matrix lead to the same estimation variance on most object pixels. For both cases, two-step multiplexing leads to an estimation variance twice as large as one-step multiplexing. Implementing a dual-detection strategy for one-step multiplexing divides the estimation variance by 2, except when the balanced-detection strategy is employed for the  $\mathbf{S}$ -matrix. For  $\mathbf{H}_1$ , the balanced detection strategy - defined in equation 3.3.11 - leads to approximately the same variance than when considering the full measurements (see section 4.4).

### 4.2.3 Proofs of the variance results ( $\mathbf{S}$ -matrix)

Here we prove the results of Table 4.2. To do so, we simply insert matrices based on  $\mathbf{S}$ -matrix into the variance expressions of Fig. 3.1.

#### One-step multiplexing

Replacing  $\mathbf{A}$  by  $\mathbf{S}$  in equation (3.1.1) and using equations (4.2.7) and (4.2.5), the estimation variance obtained for one-step multiplexing reads:

$$\begin{aligned}\mathbf{V}_S(\hat{\mathbf{x}}) &= (\mathbf{S}^{-1} \odot \mathbf{S}^{-1})\mathbf{S}\mathbf{x} \\ &= \frac{4}{(N+1)^2}\mathbf{J}_N\mathbf{S}\mathbf{x}\end{aligned}$$

i.e.

$$\boxed{\mathbf{V}_S(\hat{\mathbf{x}}) = \frac{2N}{N+1}\bar{x}\mathbf{1}_N} \quad (4.2.16)$$

Therefore, if  $N \gg 1$ , the variance reads:

$$\boxed{\mathbf{V}_S(\hat{\mathbf{x}}) \approx 2\bar{x}\mathbf{1}_N} \quad (4.2.17)$$

or, on every pixel  $i$

$$V_S(\hat{x}_i) \approx 2\bar{x} \quad \forall i \quad (4.2.18)$$

#### Two-step multiplexing

If the two multiplexing steps are based on  $\mathbf{S}$ -matrices of size  $\sqrt{N}$ , then  $\mathbf{P}^T = \mathbf{P} = \mathbf{U} = \mathbf{S}$ . Replacing in (3.2.5) and using (4.2.5) and (4.2.7) lead to:

$$\begin{aligned}\mathbf{V}_{S \otimes S}(\hat{\mathbf{x}}) &= ((\mathbf{S}^{-1} \odot \mathbf{S}^{-1})\mathbf{S}) \otimes ((\mathbf{S}^{-1} \odot \mathbf{S}^{-1})\mathbf{S})\mathbf{x} \\ &= \frac{16}{(\sqrt{N}+1)^4}(\mathbf{J}_{\sqrt{N}}\mathbf{S} \otimes \mathbf{J}_{\sqrt{N}}\mathbf{S})\mathbf{x} \\ &= \frac{16}{(\sqrt{N}+1)^4} \frac{(\sqrt{N}+1)^2}{4}(\mathbf{J}_{\sqrt{N}} \otimes \mathbf{J}_{\sqrt{N}})\mathbf{x} \\ &= \frac{4}{(\sqrt{N}+1)^2}\mathbf{J}_N\mathbf{x}\end{aligned}$$

i.e.

$$\boxed{\mathbf{V}_{S \otimes S}(\hat{\mathbf{x}}) = \frac{4N}{(\sqrt{N}+1)^2}\bar{x}\mathbf{1}_N} \quad (4.2.19)$$

Therefore, if  $N \gg 1$ , the variance reads:

$$\boxed{\mathbf{V}_{S \otimes S}(\hat{\mathbf{x}}) \approx 4\bar{x}\mathbf{1}_N} \quad (4.2.20)$$

or, on every pixel  $i$

$$V_{S \otimes S}(\hat{x}_i) \approx 4\bar{x} \quad \forall i \quad (4.2.21)$$

(We surmise that this result can be extended to  $p$ -dimensions (for  $N \gg 1$ ):  $\mathbf{V}_{S \otimes p}(\hat{\mathbf{x}}) \approx 2^p\bar{x}\mathbf{1}_{N^p}$ ).

## Dual-detection

In a dual detection scheme with a  $\mathbf{S}$ -matrix, one can define the two complementary matrices as:

$$\begin{cases} \mathbf{A}_1 = \mathbf{S} \\ \mathbf{A}_2 = \mathbf{J}_N - \mathbf{S} \end{cases} \quad (4.2.22)$$

We remind that estimation variance given in equation (3.3.10) is:

$$\mathbf{V}(\hat{\mathbf{x}}) = (\mathbf{A}^+ \odot \mathbf{A}^+) \mathbf{A} \mathbf{x}$$

with

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{bmatrix}^T$$

and

$$\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T = [(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}_1 \quad (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}_2]$$

We have:

$$\begin{aligned} \mathbf{A}^T \mathbf{A} &= \mathbf{A}_1^T \mathbf{A}_1 + \mathbf{A}_2^T \mathbf{A}_2 \\ &= \mathbf{S}^T \mathbf{S} + (\mathbf{J}_N - \mathbf{S})^T (\mathbf{J}_N - \mathbf{S}) \\ &= 2\mathbf{S}\mathbf{S} + \mathbf{J}_N \mathbf{J}_N - 2\mathbf{J}_N \mathbf{S} \\ &= 2 \frac{N+1}{4} (\mathbf{I}_N + \mathbf{J}_N) + N\mathbf{J}_N - 2 \frac{N+1}{2} \mathbf{J}_N \\ &= \frac{N+1}{2} \mathbf{I}_N + \frac{N-1}{2} \mathbf{J}_N \end{aligned}$$

Using equation (0.0.3) leads to:

$$\begin{aligned} (\mathbf{A}^T \mathbf{A})^{-1} &= \frac{2}{N+1} \mathbf{I} - \frac{\left(\frac{2}{N+1}\right)^2 \frac{N-1}{2}}{1 + \frac{2}{N+1} \frac{N-1}{2} N} \mathbf{J}_N \\ &= \frac{2}{N+1} \mathbf{I}_N - 2 \frac{N-1}{N+1} \frac{1}{N^2+1} \mathbf{J}_N \end{aligned}$$

Hence:

$$\begin{aligned} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}_1 &= \frac{2}{N+1} \mathbf{S} - 2 \frac{N-1}{N+1} \frac{1}{N^2+1} \frac{N+1}{2} \mathbf{J} \\ &= \frac{2}{N+1} \mathbf{S} - \frac{N-1}{N^2+1} \mathbf{J}_N \end{aligned}$$

and

$$\begin{aligned} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}_2 &= \frac{2}{N+1} \mathbf{J}_N - 2 \frac{N-1}{N+1} \frac{N}{N^2+1} \mathbf{J}_N - \frac{2}{N+1} \mathbf{S} + \frac{N-1}{N^2+1} \mathbf{J}_N \\ &= \frac{N+1}{N^2+1} \mathbf{J}_N - \frac{2}{N+1} \mathbf{S} \end{aligned}$$

This leads to:

$$\mathbf{A}^+ \odot \mathbf{A}^+ = \begin{bmatrix} (\mathbf{A}^T \mathbf{C})^{-1} \mathbf{A}_1 \odot (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}_1 & (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}_2 \odot (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}_2 \end{bmatrix} \quad (4.2.23)$$

$$= [\alpha_1 \mathbf{S} + \beta_1 \mathbf{J}_N \quad \alpha_2 \mathbf{S} + \beta_2 \mathbf{J}_N] \quad (4.2.24)$$

with

$$\begin{cases} \alpha_1 &= \frac{8}{(1+N)^2(1+N^2)} \\ \alpha_2 &= \frac{-8N}{(1+N)^2(1+N^2)} \\ \beta_1 &= \frac{(N-1)^2}{(1+N^2)^2} \\ \beta_2 &= \frac{(N+1)^2}{(1+N^2)^2} \end{cases} \quad (4.2.25)$$

and thus (using equation (4.2.5)):

$$\begin{aligned} (\mathbf{A}^+ \odot \mathbf{A}^+) \mathbf{A} &= \alpha_1 \mathbf{S} \mathbf{S} + \beta_1 \mathbf{J}_N \mathbf{S} + \alpha_2 \mathbf{J}_N \mathbf{S} + \beta_2 \mathbf{J}_N \mathbf{J}_N - \alpha_2 \mathbf{S} \mathbf{S} - \beta_2 \mathbf{J}_N \mathbf{S} \\ &= \frac{N+1}{4} (\alpha_1 - \alpha_2) \mathbf{I}_N + \frac{N+1}{4} (\alpha_1 + \alpha_2) \mathbf{J}_N + N \beta_2 \mathbf{J}_N + \frac{N+1}{2} (\beta_1 - \beta_2) \mathbf{J}_N \\ &= \frac{2}{N^2+1} \mathbf{I}_N + \left( \frac{2(1-N)}{(N+1)(N^2+1)} + \frac{N(N+1)(N-1)}{(N^2+1)^2} \right) \mathbf{J}_N \end{aligned}$$

Finally, we can derive the variance:  $\mathbf{V}_{Sd}(\hat{\mathbf{x}}) = (\mathbf{A}^+ \odot \mathbf{A}^+) \mathbf{A} \mathbf{x}$

$$\boxed{\mathbf{V}_{Sd}(\hat{\mathbf{x}}) = \frac{2}{N^2+1} \mathbf{x} + \left( \frac{2N(1-N)}{(N+1)(N^2+1)} + \frac{N^2(N+1)(N-1)}{(N^2+1)^2} \right) \bar{x} \mathbf{1}_N} \quad (4.2.26)$$

Therefore, if  $N \gg 1$ , the variance reads:

$$\boxed{\mathbf{V}_{Sd}(\hat{\mathbf{x}}) \approx \bar{x} \mathbf{1}_N} \quad (4.2.27)$$

Note on the balanced detection strategy:

As described in section 3.3, one could as well subtract the two measurements and thus use the equivalent matrix:

$$\mathbf{A} = \mathbf{A}_1 - \mathbf{A}_2 = 2\mathbf{S} - \mathbf{J}_N = \frac{N+1}{2} \mathbf{S}^{-1}$$

Replacing in equation (3.3.13) leads to:

$$\begin{aligned} \mathbf{V}_{Sb}(\hat{\mathbf{x}}) &= N\bar{x} \frac{4}{(N+1)^2} (\mathbf{S} \odot \mathbf{S}) \mathbf{1}_N = N\bar{x} \frac{4}{(N+1)^2} \frac{N+1}{2} \mathbf{1}_N \\ \mathbf{V}_{Sb}(\hat{\mathbf{x}}) &= 2\bar{x} \frac{N}{N+1} \mathbf{1}_N \end{aligned} \quad (4.2.28)$$

Therefore, if  $N \gg 1$ , the variance reads:

$$\boxed{\mathbf{V}_{Sb}(\hat{\mathbf{x}}) \approx 2\bar{x} \mathbf{1}_N} \quad (4.2.29)$$

or, on every pixel  $i$

$$V_{Sb}(\hat{x}_i) \approx 2\bar{x} \quad \forall i \quad (4.2.30)$$

The obtained variance with is twice higher than the variance obtained (equation (4.2.27)) with the dual detection strategy based on  $\mathbf{A} = [\mathbf{A}_1 \ \mathbf{A}_2]^T$ . Therefore, when performing dual-detection with the  $\mathbf{S}$ -matrix, it is important *not* to subtract the two measurements.

#### 4.2.4 Proofs of the variance results ( $\mathbf{H}_1$ -matrix)

Here we prove the results of Table 4.3. To do so, we simply insert matrices based on the  $\mathbf{H}_1$ - matrix into the expressions of Fig. 3.1.

## One-step multiplexing

Replacing  $\mathbf{A}$  by  $\mathbf{H}_1$  in equation (3.1.1) and using equations (4.2.8), (4.2.9) and (4.2.12), lead to the estimation variance obtained for one-step multiplexing:

$$\begin{aligned}
\mathbf{V}_{H_1}(\hat{\mathbf{x}}) &= (\mathbf{H}_1^{-1} \odot \mathbf{H}_1^{-1})\mathbf{H}_1\mathbf{x} \\
&= \left(\frac{4}{N^2}\mathbf{J}_N + \frac{N-4}{N}\mathbf{e}_1\mathbf{e}_1^T\right)\mathbf{H}_1\mathbf{x} \\
&= \left(\frac{2}{N^2}\mathbf{J}_N + \frac{N-4}{2N}\mathbf{e}_1\mathbf{e}_1^T\right)(\mathbf{H} + \mathbf{J}_N)\mathbf{x} \\
&= \left(\frac{2}{N}\mathbf{1}\mathbf{e}_1^T + \frac{N-4}{N}\mathbf{e}_1\mathbf{1}^T + \frac{2}{N}\mathbf{J}_N\right)\mathbf{x}
\end{aligned}$$

i.e.

$$\boxed{\mathbf{V}_{H_1}(\hat{\mathbf{x}}) = (2\bar{x} + \frac{2}{N}x_1)\mathbf{1}_N + (N-4)\bar{x}\mathbf{e}_1} \quad (4.2.31)$$

If  $N \gg 1$  and if  $N \gg 2x_1$  (which can often be arranged in practice), the variance reads:

$$\boxed{\mathbf{V}_{H_1}(\hat{\mathbf{x}}) \approx 2\bar{x}\mathbf{1}_N + (N-4)\bar{x}\mathbf{e}_1} \quad (4.2.32)$$

or, on pixel  $i$

$$V_{H_1}(\hat{x}_i) \approx \begin{cases} 2\bar{x} & \forall i \neq 1 \\ (N-2)\bar{x} & \text{for } i = 1 \end{cases} \quad (4.2.33)$$

On all pixels but one, the variance equals twice the object average, as with the S-multiplexing. On the first object pixel ( $i = 1$ ), the variance scales with  $(N-2)\bar{x}$ , thus the first object pixel may often be mis-estimated. The influence of the first object pixel intensity  $x_1$  is due to the structure of the Hadamard matrix with its first row and column with only ones. A similar derivation can be found in [10].

## Two-step multiplexing

If the two multiplexing steps are based on positive Hadamard-matrices of size  $\sqrt{N}$ , then  $\mathbf{U} = \mathbf{P}^T = \mathbf{P} = \mathbf{H}_1$ . Replacing in (3.2.5) and using the above result leads to:

$$\begin{aligned}
\mathbf{V}_{H_1 \otimes H_1}(\hat{\mathbf{x}}) &= (((\mathbf{H}_1^{-1} \odot \mathbf{H}_1^{-1})\mathbf{H}_1) \otimes (\mathbf{H}_1^{-1} \odot \mathbf{H}_1^{-1})\mathbf{H}_1)\mathbf{x} \\
&= \frac{1}{N}((2\mathbf{1}\mathbf{e}_1^T + (\sqrt{N}-4)\mathbf{e}_1\mathbf{1}^T + 2\mathbf{J}_{\sqrt{N}}) \otimes (2\mathbf{1}\mathbf{e}_1^T + (\sqrt{N}-4)\mathbf{e}_1\mathbf{1}^T + 2\mathbf{J}_{\sqrt{N}}))\mathbf{x} \\
&= 4\bar{x}\mathbf{1}_N + \frac{1}{N}(4\mathbf{1}\mathbf{e}_1^T \otimes \mathbf{1}\mathbf{e}_1^T + (\sqrt{N}-4)^2\mathbf{e}_1\mathbf{1}^T \otimes \mathbf{e}_1\mathbf{1}^T + 4(\mathbf{1}\mathbf{e}_1^T \otimes \mathbf{J}_{\sqrt{N}} + \mathbf{J}_{\sqrt{N}} \otimes \mathbf{1}\mathbf{e}_1^T) + \\
&2(\sqrt{N}-4)(\mathbf{1}\mathbf{e}_1^T \otimes \mathbf{e}_1\mathbf{1}^T + \mathbf{e}_1\mathbf{1}^T \otimes \mathbf{1}\mathbf{e}_1^T + \mathbf{e}_1\mathbf{1}^T \otimes \mathbf{J}_{\sqrt{N}} + \mathbf{J}_{\sqrt{N}} \otimes \mathbf{e}_1\mathbf{1}^T)\mathbf{x}
\end{aligned}$$

As expected, it results in a constant term in  $4\bar{x}$  and in many 'special pixels' given by the Kronecker products of the different elements. To isolate these special pixels it is relevant to treat the object as a 2-D object (which is the case in this two-step multiplexing scheme), as illustrated in Fig. 4.4. We consider the matrix  $\mathbf{X}$  with elements  $X_{ij}$  where  $\mathbf{x} = \text{vec}(\mathbf{X})$ . Then, the variance reshaped in 2-D reads:

$$\boxed{\mathbf{V}_{H_1 \otimes H_1}(\hat{\mathbf{X}}) = 4\bar{X} + \frac{4}{N}(X_{11} + \sum X_{1j} + \sum X_{i1})\text{vec}(\mathbf{J}_{\sqrt{N}}) + \text{vec} \begin{pmatrix} c & d & \dots & d \\ e & 0 & \dots & 0 \\ \cdot & 0 & \dots & 0 \\ e & 0 & \dots & 0 \end{pmatrix}} \quad (4.2.34)$$

with

$$\begin{cases} c &= (\sqrt{N} - 4)(\sqrt{N}\bar{X} + \frac{2}{N}(\sum X_{1j} + \sum X_{i1})) \\ d &= 2(\sqrt{N} - 4)(\bar{X} + \frac{1}{N}\sum X_{i1}) \\ e &= 2(\sqrt{N} - 4)(\bar{X} + \frac{1}{N}\sum X_{1j}) \end{cases}$$

where  $\sum X_{i1}$  is the sum of all the elements of the first column of  $\mathbf{X}$  and  $\sum X_{1j}$  is the sum of all the elements of its first row.

When  $N \gg 1$  (or more precisely, when  $\sum X_{ij} \gg \sum X_{1j}$  and  $\sum X_{ij} \gg \sum X_{i1}$ , i.e. the sum of all the object elements is much larger than the sum of the elements of its first row or column), the variance simplifies to:

$$\mathbf{V}_{H_1 \otimes H_1}(\hat{\mathbf{X}}) \approx \text{vec} \begin{pmatrix} c' & d' & \dots & d' \\ d' & f' & \dots & f' \\ \cdot & f' & \dots & f' \\ d' & f' & \dots & f' \end{pmatrix} \quad (4.2.35)$$

with

$$\begin{cases} c' &= 4\bar{X} + \sqrt{N}(\sqrt{N} - 4)\bar{X} \\ d' &= 4\bar{X} + 2(\sqrt{N} - 4)\bar{X} \\ f' &= 4\bar{X} \end{cases}$$

In 2-D, under the above assumptions, the estimation variance is equal to 4 times the object average ( $4\bar{X}$ ), on most pixels, except on its 1<sup>st</sup> line and column. There, the  $4\bar{X}$  value is supplemented by some constants ( $c', d', e'$ ) that depend on the object average  $\bar{X}$ . This specific structure is illustrated in Fig. 4.4.

### Dual-detection

In a dual detection scheme with  $\mathbf{H}_1$ , one can define the two complementary matrices as:

$$\begin{cases} \mathbf{A}_1 = \mathbf{H}_1 = \frac{1}{2}(\mathbf{J}_N + \mathbf{H}) \\ \mathbf{A}_2 = \mathbf{H}_2 = \frac{1}{2}(\mathbf{J}_N - \mathbf{H}) \end{cases} \quad (4.2.36)$$

When considering the complete measurement vector and therefore the equivalent matrix  $\mathbf{A} = [\mathbf{H}_1 \ \mathbf{H}_2]^T$ , and adapting the general calculation of section 4.4 to the Hadamard matrix, the variance reads:

$$\mathbf{V}_{Hd}(\hat{\mathbf{x}}) = \bar{x} \left(1 - \frac{N+3}{(N+1)^2}\right) \mathbf{1}_N + \frac{2}{N(N+1)} \mathbf{x} \quad (4.2.37)$$

Then, if  $N \gg 1$ :

$$\mathbf{V}_{Hd}(\hat{\mathbf{x}}) \approx \bar{x} \mathbf{1}_N \quad (4.2.38)$$

Note for the balanced detection strategy:

As described in section 3.3, one could as well subtract the two measurements and thus use the equivalent matrix:

$$\mathbf{A} = \mathbf{H}_1 - \mathbf{H}_2 = \mathbf{H} \quad (4.2.39)$$

The resulting equivalent matrix is therefore equal to the Hadamard matrix defined in equation (4.2.4). Using equation (3.3.15) it is straightforward to show that the covariance matrix reads:

$$\mathbf{\Gamma}_{H1b} = \bar{x}\mathbf{I}_N \quad (4.2.40)$$

And therefore, the estimation variance reads:

$$\mathbf{V}_{H1b}(\hat{\mathbf{x}}) = \bar{x}\mathbf{I}_N \quad (4.2.41)$$

Or, on every pixel  $i$ :

$$V_{H1b}(\hat{x}_i) = \bar{x} \quad (4.2.42)$$

Hence, if  $N \gg 1$ , simply subtracting the measurements lead to a variance  $\mathbf{V}_{H1b}$  approximately equal to the variance  $\mathbf{V}_{H1d}$  obtained when considering the full measurement vector (equation (3.3.8)):

$$\mathbf{V}_{H1b}(\hat{\mathbf{x}}) \approx \mathbf{V}_{H1d}(\hat{\mathbf{x}}) \approx \bar{x}\mathbf{I}_N \quad (4.2.43)$$

In both cases, the dual detection scheme with Hadamard matrices divides the variance by 2 as compared to one-step H1-multiplexing.

### 4.3 Positive-Cosine multiplexing

Multiplexing based on positive-cosines (or positive-sines) modulation can be implemented in many different manners (e.g. [7, 19–23]). In an analogue way to positive-Hadamard-based multiplexing, an object intensity can be modulated with positive-cosine patterns, such as in Fig. 4.2(a). Other cases of cosine-based multiplexing can be found in interferometric measurements (e.g. typical Michelson-interferometer of Fig. 4.2(b)). In both cases, the measurement, in its discrete form, can be related to the object  $\mathbf{x}$  by a general multiplexing matrix  $\mathbf{C}_1$ :

$$\mathbf{b} \sim \text{Poisson}(\mathbf{C}_1\mathbf{x}) \quad (4.3.44)$$

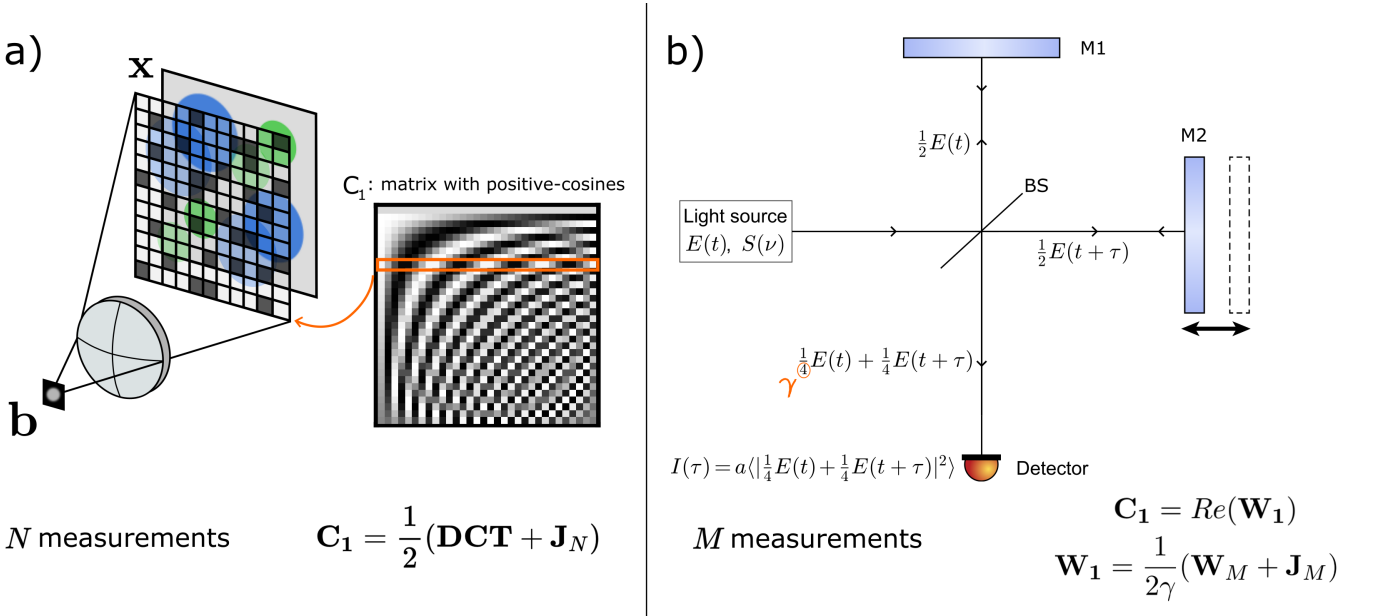


Figure 4.2: (a) Example of a positive-cosine modulation of an intensity object  $\mathbf{x}$ . (b) Scheme of a typical Michelson-interferometer. The detected intensity undergoes a modulation resulting from the interference of the non-monochromatic electric field  $E(t)$  at several delays  $\tau$ .  $I(\tau)$ : detected intensity,  $S(\nu)$ : power spectrum of  $E(t)$ , BS: beam splitter, M1: fixed mirror, M2: moving mirror inducing time delay  $\tau$  (see section 4.3.2 for more details).

### Case 1: Fig. 4.2(a)

Multiplexing based on intensity modulation is often implemented on systems similar Fig. 4.2(a). In an analogue way to positive-Hadamard-based multiplexing (section 4.2, an object intensity of size  $N$  is modulated with  $N$  positive-cosine patterns. The patterns generally derive from the discrete cosine transform (DCT) matrix, but the latter has different possible definitions [24]. The most familiar definitions of DCT are equivalent - to some normalisation factor - to the discrete Fourier transform of real numbers with even symmetry [24]. But the MSE results may depend which DCT definition is chosen. Here, we choose the multiplexing matrix  $\mathbf{C}_1$  such that its coefficients are comprised between 0 and 1:

$$\mathbf{C}_1 = \frac{1}{2}(\mathbf{DCT} + \mathbf{J}_N) \quad (4.3.45)$$

where  $\mathbf{DCT}$  is the discrete-Cosine transform matrix with coefficients comprised between -1 and +1.

### Case 2: Fig. 4.2(b)

Other types of systems, such as interferometric systems [25, 26] or [19, 27], also perform cosine-based multiplexing. For example, in the typical Michelson-interferometer of Fig. 4.2(b), the detected intensity  $I(\tau)$  undergoes a modulation, which is related by some positive-cosine transform to the the field power spectrum  $S(\nu)$  (section 4.3.2). In such systems, the matrix  $\mathbf{C}_1$  is based on the real part of the Fourier matrix, and the number of measurements is assumed to be  $M \geq N$ . In addition, there exist many different interferometers layouts and system specificities that lead to different equivalent multiplexing matrices  $\mathbf{C}_1$  [26, 28]. For such systems, we define generic positive-cosine multiplexing matrix  $\mathbf{C}_1$  related to the real part of a matrix  $\mathbf{W}_1$ :

$$\mathbf{W}_1 = \frac{1}{2\gamma}(\mathbf{W}_M + \mathbf{J}_M) \quad (4.3.46)$$

where  $\mathbf{W}_M \in \mathbb{C}_{M^2}$  is the discrete inverse Fourier transform matrix defined in equation (4.3.57) and  $\gamma$  is a real positive constant that accounts for some system specificities (for example, for the ideal Michelson interferometer of Fig. 4.2(b),  $\gamma = 4$  (section 4.3.2). Here we restrict ourselves to the case  $M \approx 2N$ .

#### 4.3.1 Results

The resulting estimation variances are shown in Table. 4.4. Note that the variances of the system of Case 1 and Case 2 cannot directly be compared since the number of measurements is not the same. For alternative definitions of the multiplexing matrix  $\mathbf{C}_1$ , the resulting variances can be derived using Table. 5.1 and Table. 5.2. Overall, the variance associated with a general positive-cosine matrix is equal to a constant which depends on the exact definition of  $\mathbf{C}_1$ .

Importantly, we note that for Case 1, the variance is twice larger than for positive-Hadamard multiplexing with  $\mathbf{H}_1$  (Table. 4.3). This may be surprising but can be explained because  $\mathbf{H}_1$  and  $\mathbf{C}_1$  - as defined in equation (4.2.8) and (4.3.45), respectively - do not have the same norm <sup>2</sup>. This means that, for an identical system, the energy transmitted by  $\mathbf{H}_1$  is not the same as the energy transmitted by  $\mathbf{C}_1$ . A matrix  $\mathbf{C}_1$  defined such that its norm is equal to the norm of  $\mathbf{H}_1$  leads to the same variance as with  $\mathbf{H}_1$ , but would comprise some negative coefficients.

For Case 2, the proofs are provided in section 4.3.2. From these proofs, the variance for Case 1 is deduced, and is verified numerically in Fig. S2 (Supp Methods).

---

<sup>2</sup>largest singular value

|  | One-step multiplexing                   | Two-step multiplexing                       | Dual-detection                         |                               |
|--|---|---|--|-------------------------------|
| <b>A</b> -matrix   | $\mathbf{C}_1$                          | $\mathbf{C}_1 \otimes \mathbf{C}_1$         | $[\mathbf{C}_1 \ \mathbf{C}_2]^T$      | $\mathbf{C}_1 - \mathbf{C}_2$ |
| General positive-cosine matrix $\mathbf{C}_1$  |   |   |  |                               |
| $V(\hat{x}_i)$   | $k'\bar{x}$                             | $k'\bar{x}$                                 | $k''\bar{x}$                           |                               |
| Case 1: $\mathbf{C}_1$ of size $N \times N$ , with coefficients between 0 and 1. $\mathbf{C}_1 = \frac{1}{2}(\mathbf{DCT} + \mathbf{J}_N)$ |   |   |  |                               |
| $V(\hat{x}_i)$   | $4\bar{x}$                              | $16\bar{x}$                                 | $2\bar{x}$                             |                               |
| Case 2: $\mathbf{C}_1$ of size $M \times M$ , related to real part of $\mathbf{W}_1 = \frac{1}{2\gamma}(\mathbf{W}_M + \mathbf{J}_M)$      |   |   |  |                               |
| $V(\hat{x}_i)$   | $2\gamma\bar{x}$ ( $\forall i \neq 1$ ) | $4\gamma^2\bar{x}$ ( $\forall i \neq n_1$ ) | $\gamma\bar{x}$ ( $\forall i \neq 1$ ) |                               |
|  | $2\gamma(M-2)\bar{x}$ ( $i=1$ )         | $c', d'$ ( $\forall i \neq n_1$ )           | $2\gamma\bar{x}$ ( $i=1$ )             |                               |

Table 4.4: Multiplexing-schemes and associated estimation variances in the three multiplexing modalities. The variance expressions on pixel  $i$  are valid for large number of pixels  $N \gg 1$  (exact expressions in the proofs). The results in black are valid on most pixels, while the results in gray are valid only on specific pixels.  $k'$ ,  $k''$  and  $k'''$  are constants that depend on the exact definition of  $\mathbf{C}_1$ .  $\gamma \in \mathbb{R}_+^*$  is a constant. **A**-matrix: Global multiplexing matrix.  $\mathbf{C}_2$  is the complementary matrix of  $\mathbf{C}_1$ . The last column corresponds to a balanced-detection strategy (section 3.3.11).  $\bar{x}$ : object intensity average,  $N$ : object size,  $M = 2(N-1)$ ,  $n_1$ : specific pixels on the first row and column of the image,  $c', d'$  are given in equation (4.3.75).

**Note:** For positive-cosine multiplexing, common alternative solutions to the balanced-detection scheme presented here exist (e.g. the four-step phase-shifting method), with the aim of removing the detected DC component [7, 22]. Such strategies can improve the MSE (at the expense of a higher number of measurements), but are not considered in this work.

### 4.3.2 Proofs for Case 2 (Fig. 4.2(b))

#### Model: positive-multiplexing via interferometric measurements

First, we consider the physical model Michelson-based interferometric measurements. We assume that we seek to estimate a power spectrum  $S(\nu)$  from intensity measurements  $I(\tau)$  at different time delays  $\tau$ . Here, the constant  $\gamma$  may depend on the specific method of modulation and detection [26, 28], for instance on the interferometer design, on the beam-splitter, etc. To a first approximation (perfect beam-splitter, no apodization, etc), the measured intensity  $I(\tau)$  by the above system reads:

$$I(\tau) = \int_{-\infty}^{\infty} \left| \frac{1}{\sqrt{4\gamma}}E(t) + \frac{1}{\sqrt{4\gamma}}E(t+\tau) \right|^2 dt = \frac{1}{2\gamma}(I_0 + R(\tau)) \quad (4.3.47)$$

where  $I_0$  is the intensity of the initial field  $E(t)$  :

$$I_0 = \int_{-\infty}^{\infty} |E(t)|^2 dt = \int_{-\infty}^{\infty} S(\nu) d\nu \quad (4.3.48)$$

and  $R(\tau)$  is the field autocorrelation, i.e. the inverse Fourier-transform of the power spectrum  $S(\nu)$ :

$$R(\tau) = \int_{-\infty}^{\infty} E(t)E(t+\tau) dt = \int_{-\infty}^{\infty} S(\nu)e^{2i\pi\nu\tau} d\nu \quad (4.3.49)$$



Therefore, the detected intensity becomes:

$$I(\tau) = \frac{1}{2\gamma} (I_0 + \int_{-\infty}^{\infty} e^{2i\pi\nu\tau} S(\nu) d\nu) = \frac{1}{2\gamma} \int_{-\infty}^{\infty} (1 + e^{2i\pi\nu\tau}) S(\nu) d\nu \quad (4.3.50)$$

The least-square estimation of the power spectrum is thus given by the real part of the Fourier-transform:

$$\hat{S}(\nu) = \text{Re} \left( \int_{-\infty}^{\infty} (2\gamma I(\tau) - I_0) e^{-2i\pi\nu\tau} d\tau \right) \quad (4.3.51)$$

Since  $2\gamma I(\tau) - I_0 \in \mathbb{R}$  and  $\hat{S}(\nu)$  is even, the physical power spectrum is thus obtained by selecting only the positive frequencies, and multiplying the intensity by 2 in order to fulfil (4.3.48). Note that then, the number of spanned delays  $\tau$  should be at least twice higher than the number of spanned frequencies  $\nu$ .

### Associated discrete model

The above model can be written in a discrete form using the following variables:

- $\mathbf{b} \in \mathbb{N}_M^+$ : photon counts of the detected intensity  $I(\tau_m)$  with  $\tau_m = 0 : M - 1$
- $\mathbf{W}_M \in \mathbb{C}_{M^2}$ : discrete inverse Fourier transform matrix, defined in equation (4.3.57).
- $\mathbf{x} \in \mathbb{R}_N^+$ : discrete power spectrum  $S(\nu_n)$  with  $\nu_n \in [0 : N - 1]$
- $\mathbf{x}^s \in \mathbb{R}_M^+$ : symmetrized discrete power spectrum.

Therefore, we consider the following model:

$$\mathbf{b} \sim \text{Poisson}(\mathbf{W}_1 \mathbf{x}^s) \quad (4.3.52)$$

where  $\mathbf{W}_1$  is defined as in equation 4.3.46  $\mathbf{b}$  is the measurement vector of size  $M = 2(N - 1)$ .  $\mathbf{x}_s \in \mathbb{R}_M^+$  is a temporary variable built from  $\mathbf{x}$  and its even-symmetric:

$$\mathbf{x}^s = \frac{1}{2} [x_0, x_1, \dots, x_{N-1}, x_{N-2}, \dots, x_1]^T. \quad (4.3.53)$$

Then, even if  $\mathbf{W}_1$  is a complex quantity,  $\mathbf{W}_1 \mathbf{x}_s$  is a real positive quantity<sup>3</sup> and so the measurement  $\mathbf{b}$  can be defined ( $\mathbf{b} \in \mathbb{N}^N$ ). Note that we define  $\mathbf{x}^s$  so that  $\mathbf{x}$  and  $\mathbf{x}^s$  have the same energy:

$$I_0 = \sum x_n = \sum x_n^s = N\bar{\mathbf{x}} = 2(N - 1)\bar{\mathbf{x}}^s \quad (4.3.54)$$

The LS-estimate of  $\mathbf{x}^s$  is:

$$\hat{\mathbf{x}}^s = \mathbf{W}_1^{-1} \mathbf{b} = \frac{1}{2} [\hat{x}_0 \dots \hat{x}_{N-1}, \hat{x}_{N-2}, \dots, \hat{x}_1]^T \quad (4.3.55)$$

The final estimate, denoted  $\hat{\mathbf{x}}_{Re}$  is obtained by taking the real part of  $\hat{\mathbf{x}}^s$ , selecting half of it and multiplying it by 2 (according to equation (4.3.53))<sup>4</sup>

$$\hat{\mathbf{x}}_{Re} = \text{Re}(\hat{\mathbf{x}}) = 2\text{Re}(\hat{\mathbf{x}}^s)_{[0:N-1]} = \text{Re}[\hat{x}_0, \dots, \hat{x}_{N-1}]^T \quad (4.3.56)$$

<sup>3</sup>If  $g$  is real and even, its Fourier-transform is real and even.

<sup>4</sup>Since  $\mathbf{b}$  is real,  $\text{Re}(\hat{\mathbf{x}}^s)$  is even, therefore this is equivalent to averaging both parts of  $\text{Re}(\hat{\mathbf{x}}^s)$ .

## The discrete Fourier transform (DFT) matrix

$\mathbf{W}_M \in \mathbb{C}_{M^2}$  is the inverse of the discrete Fourier transform matrix. Its  $(k, m)$ -element reads, for  $k, m = (0, 1, \dots, M-1)$ :

$$w_M^{km} = e^{\frac{2i\pi km}{M}} \quad (4.3.57)$$

$$\mathbf{W}_M = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w_M & w_M^2 & \cdots & w_M^{M-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w_M^{M-1} & w_M^{2(M-1)} & \cdots & w_M^{(M-1)^2} \end{pmatrix}$$

It is a symmetric matrix ( $\mathbf{W}_M^T = \mathbf{W}_M$ ) with properties similar to the Hadamard matrix  $\mathbf{H}$ :

$$\mathbf{W}_M \mathbf{W}_M^* = M \mathbf{I}_M \quad (4.3.58)$$

$$\mathbf{W}_M \odot \mathbf{W}_M^* = \mathbf{J}_M \quad (4.3.59)$$

$$\mathbf{J}_N \mathbf{W}_N = \mathbf{J}_N \mathbf{W}_N^* = N \mathbf{e}_1 \mathbf{e}_1^T \quad (4.3.60)$$

$$\mathbf{W}_N \mathbf{J}_N = \mathbf{W}_N^* \mathbf{J}_N = N \mathbf{e}_1 \mathbf{1}^T \quad (4.3.61)$$

$\mathbf{W}_1$  is the complex analogous of the positive Hadamard matrix  $\mathbf{H}_1$ . Using (0.0.3), its inverse reads:

$$\mathbf{W}_1^{-1} = 2\gamma(\mathbf{W}_M + \mathbf{J}_M)^{-1} = \frac{2\gamma}{M}(\mathbf{W}_M^* - \frac{M}{2}\mathbf{e}_1 \mathbf{e}_1^T) \quad (4.3.62)$$

In addition, since  $\mathbf{W}_M \in \mathbb{C}_{M^2}$ , we need to consider the estimation variance in the complex case. Using the complex definition of the covariance ( $\mathbf{\Gamma} = \langle \delta \hat{\mathbf{x}} \delta \hat{\mathbf{x}}^H \rangle$ ) and equation (0.0.2), the estimation variance reads:

$$\mathbf{V}(\hat{\mathbf{x}}) = (\mathbf{A}^{-1} \odot (\mathbf{A}^{-1})^*) \mathbf{A} \mathbf{x} \quad (4.3.63)$$

The variance of the estimate real part  $\hat{\mathbf{x}}_{Re} = Re(\hat{\mathbf{x}})$  is:

$$\mathbf{V}(\hat{\mathbf{x}}_{Re}) = \frac{1}{2} Re(diag(\langle \delta \hat{\mathbf{x}} \delta \hat{\mathbf{x}}^H \rangle + \langle \delta \hat{\mathbf{x}} \delta \hat{\mathbf{x}}^T \rangle)) = \frac{1}{2} Re((\mathbf{A}^{-1} \odot (\mathbf{A}^{-1})^* + \mathbf{A}^{-1} \odot \mathbf{A}^{-1}) \mathbf{A} \mathbf{x}) \quad (4.3.64)$$

## One-step multiplexing

To calculate the variance of the real part of  $\hat{\mathbf{x}}$ , we first calculate the variance of the real part of the symmetrized object  $\hat{\mathbf{x}}^s$ . It is obtained by inserting  $\mathbf{W}_1^{-1}$  into (4.3.64) :

$$\begin{aligned} \mathbf{V}_{W1}(\hat{\mathbf{x}}_{Re}^s) &= \frac{2\gamma}{2M^2} Re \left( \left( (\mathbf{W}_M^* - \frac{M}{2}\mathbf{e}_1 \mathbf{e}_1^T) \odot (\mathbf{W}_M - \frac{M}{2}\mathbf{e}_1 \mathbf{e}_1^T) + (\mathbf{W}_M^* - \frac{M}{2}\mathbf{e}_1 \mathbf{e}_1^T)^{\odot 2} \right) (\mathbf{W}_M + \mathbf{J}_M) \mathbf{x}^s \right) \\ &= \frac{2\gamma}{2M^2} Re \left( (\mathbf{J}_M + \mathbf{W}_M^* \odot \mathbf{W}_M^* + \frac{M(M-4)}{2}\mathbf{e}_1 \mathbf{e}_1^T) (\mathbf{W}_M + \mathbf{J}_M) \mathbf{x}^s \right) \\ &= \frac{2\gamma}{2M^2} Re \left( (M\mathbf{J}_M + \mathbf{J}_M \mathbf{W}_M + \widetilde{\mathbf{W}}_M^* \mathbf{W}_M + \widetilde{\mathbf{W}}_M^* \mathbf{J}_M + M^2(M-4)\mathbf{e}_1 \mathbf{1}^T) \mathbf{x}^s \right) \end{aligned}$$

where we used the fact that  $\mathbf{e}_1 \mathbf{e}_1^T \mathbf{W}_M = \mathbf{e}_1 \mathbf{e}_1^T \mathbf{J}_M = \mathbf{e}_1 \mathbf{1}_M^T$ , and defined  $\widetilde{\mathbf{W}}_M^* = \mathbf{W}_M^* \odot \mathbf{W}_M^*$ . Each matrix element of  $\widetilde{\mathbf{W}}_M^* \mathbf{W}_M$  is:

$$r_{km} = \sum_{h=0}^{M-1} (e^{\frac{2i\pi}{M}(m-2k)h})^h = \begin{cases} M & \text{if } m - 2k = 0(M) \\ 0 & \text{otherwise.} \end{cases} \quad (4.3.65)$$

In other words, the matrix  $\widetilde{\mathbf{W}}_M^* \mathbf{W}_M$  selects only the even pixels of  $\mathbf{x}^s$  in a symmetric manner. Therefore, we can write

$$(\widetilde{\mathbf{W}}_M^* \mathbf{W}_M) \mathbf{x}^s = M \begin{bmatrix} \mathbf{v} \\ \mathbf{v} \end{bmatrix} \quad (4.3.66)$$

where

$$\mathbf{v} = (x_0^s, x_2^s, \dots, x_{M-2}^s, x_0^s, x_2^s, \dots, x_{M-2}^s)^T$$

In addition, using (4.3.61) and noting that all the terms are real quantities, leads to:

$$\mathbf{V}_{W1}(\hat{\mathbf{x}}_{Re}^s) = \gamma \bar{x}^s \left( \mathbf{1}_M + \begin{bmatrix} \mathbf{e}_{1,N} \\ \mathbf{e}_{1,N} \end{bmatrix} + (M-4) \begin{bmatrix} \mathbf{e}_{1,N} \\ \mathbf{0}_N \end{bmatrix} \right) + \frac{\gamma}{M} \left( x_0^s \mathbf{1}_M + \begin{bmatrix} \mathbf{v} \\ \mathbf{v} \end{bmatrix} \right) \quad (4.3.67)$$

Then,  $\hat{\mathbf{x}}_{Re}$  is obtained from  $\hat{\mathbf{x}}_{Re}^s$  with equation (4.3.56). Further applying  $\bar{x}^s \approx \bar{x}/2$  (equation (4.3.54), if  $N \gg 1$ ) leads to an overall variance of:

$$\mathbf{V}_{W1}(\hat{\mathbf{x}}_{Re}) = 2\gamma \bar{x} (\mathbf{1}_N + (M-3)\mathbf{e}_{1,N}) + \frac{2\gamma}{M} (x_0 \mathbf{1}_N + 2\mathbf{v}) \quad (4.3.68)$$

Therefore, if  $M$  is large compared to the components of vector  $\mathbf{v}$ :

$$\mathbf{V}_{W1}(\hat{\mathbf{x}}_{Re}) \approx 2\gamma \bar{x} (\mathbf{1}_N + (M-3)\mathbf{e}_{1,N}) \quad (4.3.69)$$

i.e. on pixel  $i$ :

$$V_{W1}(\hat{x}_{i,Re}) \approx \begin{cases} 2\gamma \bar{x} & \forall i \neq 0 \\ 2\gamma(M-2)\bar{x} & \text{for } i = 0 \end{cases} \quad (4.3.70)$$

Note that an error is also present in the imaginary part of the estimate  $\hat{\mathbf{x}}^s$ . A similar but simpler calculation shows that the variance of the complex estimate  $\hat{\mathbf{x}}^s$  is twice higher than in the above expressions: the variance seems to spread equally in the real and imaginary part of the estimate.

Note also that if the total energy  $I_0 = \sum \mathbf{x}$  is known *a priori*, the large variance of the first pixel can be attenuated: the estimation variance is the same as above, except for the term proportional to  $(M-2)\bar{x}\mathbf{e}_1$ . Indeed, in this case, the model reads:

$$\mathbf{b} \sim \text{Poisson}\left(\frac{1}{2\gamma}(I_0 \mathbf{1}_M + \mathbf{W}_M \mathbf{x}^s)\right) \quad (4.3.71)$$

The estimate is:

$$\hat{\mathbf{x}}^s = \mathbf{W}_M^{-1}(2\gamma \mathbf{b} - I_0 \mathbf{1}_M) \quad (4.3.72)$$

Using the same kind of calculations as above, when if  $M$  is large compared to the components of vector  $\mathbf{v}$ , the variance of the estimate real part reads:

$$\mathbf{V}(\hat{\mathbf{x}}_{Re}) \approx 2\gamma \bar{x} \mathbf{1}_N$$

## Two-step multiplexing

In the case of two-step multiplexing with positive Fourier-matrices of size  $\sqrt{N}$ , the estimation variance is obtained by replacing  $\mathbf{U}$  and  $\mathbf{P}$  by  $\mathbf{W}_1$  in (3.2.5). Using (4.3.64), the general formula for the variance reads:

$$\mathbf{V}_{W1 \otimes W1}(\hat{\mathbf{x}}_{Re}^s) = \frac{1}{2} \text{Re} \left( ((\mathbf{W}_1^{-1} \odot (\mathbf{W}_1^{-1})^*) \mathbf{W}_1)^{\otimes 2} + ((\mathbf{W}_1^{-1} \odot \mathbf{W}_1^{-1}) \mathbf{W}_1)^{\otimes 2} \right) \quad (4.3.73)$$

Inserting (4.3.62) and performing a similar calculation that in section 4.2.4 with the 2D-object  $\mathbf{X}$  defined by  $\mathbf{x} = \text{vec}(\mathbf{X})$  leads to:

$$\mathbf{V}_{W_1 \otimes W_1}(\hat{\mathbf{x}}_{Re}) = \frac{4\gamma^2}{4} \left( 4\bar{X} + \frac{4}{M}(X_{00} + \sum X_{0j} + \sum X_{i0}) \text{vec}(\mathbf{J}_{\sqrt{N}}) + \text{vec} \begin{pmatrix} c & d_1 & \cdots & d_{N-1} \\ e_1 & f_{11} & \cdots & f_{1(N-1)} \\ \vdots & f_{21} & \cdots & f_{2(N-1)} \\ e_{N-1} & f_{(N-1)1} & \cdots & f_{(N-1)(N-1)} \end{pmatrix} \right) \quad (4.3.74)$$

where

$$\begin{cases} c &= 2(M - 4\sqrt{M} + 4)\bar{X} + \frac{4}{M}(\sqrt{M} - 3)(\sum X_{0j} + \sum X_{i0}) + \frac{4}{M}X_{00} \\ d_j &= 2(\sqrt{M} - 4)\bar{X} + \frac{2}{M}((\sqrt{M} - 4) \sum X_{i0} + 2(\sqrt{M} - 2) \sum X_{0(2j)}^s) + \frac{8}{M}X_{0(2j)}^s \\ e_i &= 2(\sqrt{M} - 4)\bar{X} + \frac{2}{M}((\sqrt{M} - 4) \sum X_{0j} + 2(\sqrt{M} - 2) \sum X_{(2i)0}^s) + \frac{8}{M}X_{(2i)0}^s \\ f_{ij} &= \frac{8}{M}X_{(2i)(2j)}^s \end{cases}$$

As expected, the variance results in a constant term in  $\gamma^2\bar{X}$  and in many 'special pixels' given by the Kronecker products of the different elements.  $\sum X_{i0}$  is the sum of all the elements of the  $0^{th}$  column of  $\mathbf{X}$ ,  $\sum X_{0j}$  is the sum of all the elements of its  $0^{th}$  row,  $X_{(2i)(2j)}^s$  is an element of a even row and column of the symmetrised object  $\mathbf{X}^s$ . When  $M \gg 1$ , the variance simplifies to:

$$\mathbf{V}_{W_1 \otimes W_1}(\hat{\mathbf{x}}_{Re}) \approx \text{vec} \begin{pmatrix} c' & d' & \dots & d' \\ d' & f' & \dots & f' \\ \cdot & f' & \dots & f' \\ d' & f' & \dots & f' \end{pmatrix} \quad (4.3.75)$$

where

$$\begin{cases} c' &= 4\gamma^2\bar{X} + 2\gamma^2(M - 4\sqrt{M} + 4)\bar{X} \\ d' &= 4\gamma^2\bar{X} + 2\gamma^2(\sqrt{M} - 4)\bar{X} \\ f' &= 4\gamma^2\bar{X} \end{cases}$$

The 2-D image of the above variance is thus approximately equal to  $\gamma^2\bar{X}$  on most pixels, except on its  $1^{st}$  row,  $1^{st}$  column, and on some specific pixels which influence is negligible if  $M \gg 1$  (see for example Fig. 4.4).

## Dual detection

In positive-cosine multiplexing, many different strategies may be implemented to remove the DC component, such as four-step-shifting detection [7, 22]. Here, we restrict ourselves to the dual detection as defined in section 3.3. In a dual detection scheme with  $\mathbf{W}_1$ , one can define the two complementary

matrices as:

$$\begin{cases} \mathbf{A}_1 = \mathbf{W}_1 = \frac{1}{2\gamma}(\mathbf{J}_M + \mathbf{W}_M) \\ \mathbf{A}_2 = \mathbf{W}_2 = \frac{1}{2\gamma}(\mathbf{J}_M - \mathbf{W}_M) \end{cases} \quad (4.3.76)$$

The two dual measurements are therefore considered to be  $\pi$ -phase-shifted. In section 4.4, we show that for the dual detection scheme as defined above (equation (4.3.76)), using the equivalent matrix  $\mathbf{A} = [\mathbf{A}_1 \ \mathbf{A}_2]^T$  or the equivalent matrix  $\mathbf{A} = \mathbf{A}_1 - \mathbf{A}_2$  leads to approximately the same estimation variance when  $M \gg 1$ . In other words, if  $M \gg 1$ , simply subtracting the measurements lead to a variance  $\mathbf{V}_{W1b}$  approximately equal to the variance  $\mathbf{V}_{W1d}$  obtained when considering the full measurement vector (equation (3.3.8)):

$$\mathbf{V}_{W1b}(\hat{\mathbf{x}}) \approx \mathbf{V}_{W1d}(\hat{\mathbf{x}}) \quad (4.3.77)$$

Below, we derive the estimation variance for the balanced-detection strategy, since it easier to derive, and since this strategy presents a substantial advantage of computational efficiency while preserving the SNR.

In the balanced-detection strategy, the following relations are verified:

$$\begin{cases} \mathbf{A} = \mathbf{A}_1 - \mathbf{A}_2 = \frac{1}{\gamma}\mathbf{W}_M \\ \mathbf{A}_1 + \mathbf{A}_2 = \frac{1}{\gamma}\mathbf{J}_M \end{cases} \quad (4.3.78)$$

Using equations (3.3.13), (4.3.64) and (4.3.59) leads to the estimation variance:

$$\begin{aligned} \mathbf{V}_{W1b}(\hat{\mathbf{x}}_{Re}^s) &= \frac{\gamma}{2M} \bar{x}^s Re \left( (\mathbf{W}_M^* \odot \mathbf{W}_M + \mathbf{W}_M^* \odot \mathbf{W}_M^*) \mathbf{1}_M \right) \\ &= \frac{\gamma}{2M} \bar{x}^s Re \left( (\mathbf{J}_M + \widetilde{\mathbf{W}}_M^*) \mathbf{1}_M \right) \\ &= \frac{\gamma}{2} \bar{x}^s \mathbf{1}_M + \frac{\gamma}{2} \bar{x}^s \begin{bmatrix} \mathbf{e}_{1,N} \\ \mathbf{e}_{1,N} \end{bmatrix} \end{aligned}$$

where  $\widetilde{\mathbf{W}}_M^* = \mathbf{W}_M^* \odot \mathbf{W}_M^*$ . The variance of the real part of the non-symmetric estimate  $\hat{\mathbf{x}}$  then reads:

$$\boxed{\mathbf{V}_{W1b}(\hat{\mathbf{x}}_{Re}) = \gamma \bar{x} (\mathbf{1}_N + \mathbf{e}_{1,N})} \quad (4.3.79)$$

Therefore,

$$V_{W1b}(\hat{x}_i) = \begin{cases} \gamma \bar{x} & \forall i \neq 0 \\ 2\gamma \bar{x} & \text{for } i = 0 \end{cases} \quad (4.3.80)$$

### Model with cosines-expressions

Since  $S(\nu) \in \mathbb{R}$  and  $R(\tau)$  is hermitian, the detected intensity can also be expressed as:

$$I(\tau) = \frac{1}{2\gamma} (I_0 + \int_{-\infty}^{\infty} S(\nu) \cos(2\pi\nu\tau) d\nu) \quad (4.3.81)$$

and the least-square estimation of the power spectrum is thus given by:

$$\hat{S}(\nu) = \int_{-\infty}^{\infty} (2\gamma I(\tau) - I_0) \cos(2\pi\nu\tau) d\tau \quad (4.3.82)$$

In this case, the discrete model is obtained by replacing the discrete Fourier matrix by a discrete cosine matrix of size  $M \times N$ , for example  $\mathbf{C}_{M,N}$  with elements  $c_{mn} = \cos(2\pi \frac{nm}{M})$ . Then, (4.3.81) becomes:

$$\mathbf{b} \sim \text{Poisson}\left(\frac{1}{2\gamma}(I_0 \mathbf{1}_M + \mathbf{C}_{M,N} \mathbf{x})\right) \quad (4.3.83)$$

and (4.3.82):

$$\hat{\mathbf{x}} = \mathbf{C}_{M,N}^+(2\gamma \mathbf{b} - I_0 \mathbf{1}_M) \quad (4.3.84)$$

### 4.3.3 Proofs for Case 1 (Fig. 4.2(a))

The same kind of proof could be derived for the matrix  $\mathbf{C}_1$  defined in equation 4.3.45. We do not provide a full proof. We simply note that since the number of measurements is twice smaller than for Case 2, the variance is expected to be twice larger than for Case 2 with  $\gamma = 1$ . This is verified numerically in Fig. 4.4 and Fig. S2.

## 4.4 Dual detection: Justification for the balanced-strategy

In this section we show that, for dual detection (section 3.3), for a subset of matrices  $\mathbf{A}$ , which include the Hadamard matrix  $\mathbf{H}$  and the Fourier matrix  $\mathbf{W}$ , the estimation variance associated to the balanced detection strategy ( $\mathbf{V}_{Ab}$ ) is approximately equal to the variance obtained when taking all the measurements ( $\mathbf{V}_{Ad}$ ). This is important because in practise, it is the balanced detection strategy which is often used. For both Hadamard and Fourier-based multiplexing, the dual detection scheme is defined as:

$$\left\{ \begin{array}{l} \mathbf{A}_1 + \mathbf{A}_2 = \mathbf{J}_N \\ \mathbf{A}_1 - \mathbf{A}_2 = \mathbf{A} \\ \mathbf{A}_1 = \frac{1}{2}(\mathbf{J}_N + \mathbf{A}) \\ \mathbf{A}_2 = \frac{1}{2}(\mathbf{J}_N - \mathbf{A}) \end{array} \right. \quad (4.4.85)$$

with  $\mathbf{A} = \mathbf{H}$  or  $\mathbf{A} = \mathbf{W}$ .  $\mathbf{A} \in \mathbb{C}^{N^2}$ .

In the following, we show that, if  $N \gg 1$ :

$$\left\{ \begin{array}{l} |a_{ij}| = 1 \\ \mathbf{A}\mathbf{A}^H = N\mathbf{I}_N \\ \mathbf{J}_N \mathbf{A}^H = N\mathbf{1}\mathbf{e}_1^T \end{array} \right. \implies \mathbf{V}_{Ad}(\hat{\mathbf{x}}) \approx \mathbf{V}_{Ab}(\hat{\mathbf{x}}) \approx \bar{x}\mathbf{1}_N \quad (4.4.86)$$

Therefore, if the matrix  $\mathbf{A} \in \mathbb{C}^{N \times N}$  respects the above conditions <sup>5</sup> (modulus of each entry is unity, matrix rows pairwise orthogonal, sum of all rows is zero except for the first row), both dual detections strategies lead to the same estimation variance provided  $N \gg 1$ . This class of matrices include the Fourier matrix  $\mathbf{W}$  and Hadamard matrix  $\mathbf{H}$ . Therefore, for Hadamard-based or Fourier-based positive dual-multiplexing, it is relevant to employ a balanced strategy: it is relevant to simply subtract the measurements vectors instead of considering the more complicated full model.

Proof:

When the full measurements are considered, the equivalent multiplexing matrix is:

$$\mathbf{C} = [\mathbf{A}_1 \ \mathbf{A}_2]^T \quad (4.4.87)$$

---

<sup>5</sup>A matrix  $\mathbf{A} \in \mathbb{C}^{N \times N}$  fulfilling the two first conditions is called a *complex Hadamard matrix* [29, 30]

and the estimation variance in the complex case (adapted from equation (3.3.10)):

$$\mathbf{V}_{Ad}(\hat{\mathbf{x}}) = (\mathbf{C}^+ \odot (\mathbf{C}^+)^*) \mathbf{C} \mathbf{x} \quad (4.4.88)$$

( $\mathbf{C}^H \mathbf{C}$  is supposed not singular). Using the conditions of (4.4.86) and equation (0.0.3) leads to:

$$\begin{aligned} \mathbf{C}^+ &= \left[ \mathbf{A}_1^H \mathbf{A}_1 + \mathbf{A}_2^H \mathbf{A}_2 \right]^{-1} \mathbf{C}^H \\ &= 2(\mathbf{A}^H \mathbf{A} + N \mathbf{J}_N)^{-1} \mathbf{C}^H \\ &= \frac{2}{N} \left( \mathbf{I}_N - \frac{\mathbf{J}_N}{N+1} \mathbf{C}^H \right) \\ &= \frac{2}{N} \left( \mathbf{I}_N - \frac{\mathbf{J}_N}{N+1} \right) \begin{bmatrix} \mathbf{A}_1^H & \mathbf{A}_2^H \end{bmatrix} \\ &= \frac{1}{N} \begin{bmatrix} \frac{1}{N+1} \mathbf{J}_N + \mathbf{A}^H - \frac{1}{N+1} \mathbf{J}_N \mathbf{A}^H & \frac{1}{N+1} \mathbf{J}_N - \mathbf{A}^H + \frac{1}{N+1} \mathbf{J}_N \mathbf{A}^H \end{bmatrix} \\ &= \frac{1}{N} \begin{bmatrix} \frac{1}{N+1} \mathbf{J}_N + \mathbf{A}^H - \frac{N}{N+1} \mathbf{1e}_1^T & \frac{1}{N+1} \mathbf{J}_N - \mathbf{A}^H + \frac{N}{N+1} \mathbf{1e}_1^T \end{bmatrix} \end{aligned}$$

Replacing in (3.3.10) leads to:

$$\begin{aligned} \mathbf{V}_{Ad} &= \frac{1}{N^2(N+1)} \begin{bmatrix} \frac{1+(N+1)^2}{(N+1)} \mathbf{J}_N + \mathbf{A}^H + \mathbf{A}^T - \frac{N(N+4)}{N+1} \mathbf{1e}_1^T & \frac{1+(N+1)^2}{(N+1)} \mathbf{J}_N - \mathbf{A}^H + \mathbf{A}^T - \frac{N^2}{N+1} \mathbf{1e}_1^T \end{bmatrix} \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{x} \\ &= \frac{1}{N^2} \left( N \left( 1 - \frac{N+3}{(N+1)^2} \right) \mathbf{J}_N + \frac{1}{N+1} (\mathbf{A}^H + \mathbf{A}^T) \mathbf{A} \right) \mathbf{x} \\ &= \bar{x} \left( 1 - \frac{N+3}{(N+1)^2} \right) \mathbf{1}_N + \frac{1}{N(N+1)} \left( 1 + \frac{1}{N} \mathbf{A}^T \mathbf{A} \right) \mathbf{x} \end{aligned}$$

Therefore, if  $N \gg 1$ , and since  $|a_{ij}| = 1$ , the variance obtained with the complete measurement reads:

$$\mathbf{V}_{Ad}(\hat{\mathbf{x}}) \approx \bar{x} \mathbf{1}_N \quad (4.4.89)$$

On the other hand, when subtracting the measurements, the equivalent multiplexing matrix is

$$\mathbf{A}_1 - \mathbf{A}_2 = \mathbf{A} \quad (4.4.90)$$

and the associated estimation variance is (adapted from equation (3.3.13)) is:

$$\mathbf{V}_{Ab}(\hat{\mathbf{x}}) = N \bar{x} (\mathbf{A}^{-1} \odot (\mathbf{A}^{-1})^*) \mathbf{1}_N \quad (4.4.91)$$

Then, using the properties on the left of equation (4.4.86), the variance reads:

$$\mathbf{V}_{Ab}(\hat{\mathbf{x}}) = N \bar{x} \frac{1}{N^2} \mathbf{J}_N \mathbf{1}_N = \bar{x} \mathbf{1}_N \quad (4.4.92)$$

## 4.5 Numerical results

Here, we support and illustrate the theoretical proofs with some simulations in 1-D and 2-D.

# 1-D Multiplexing

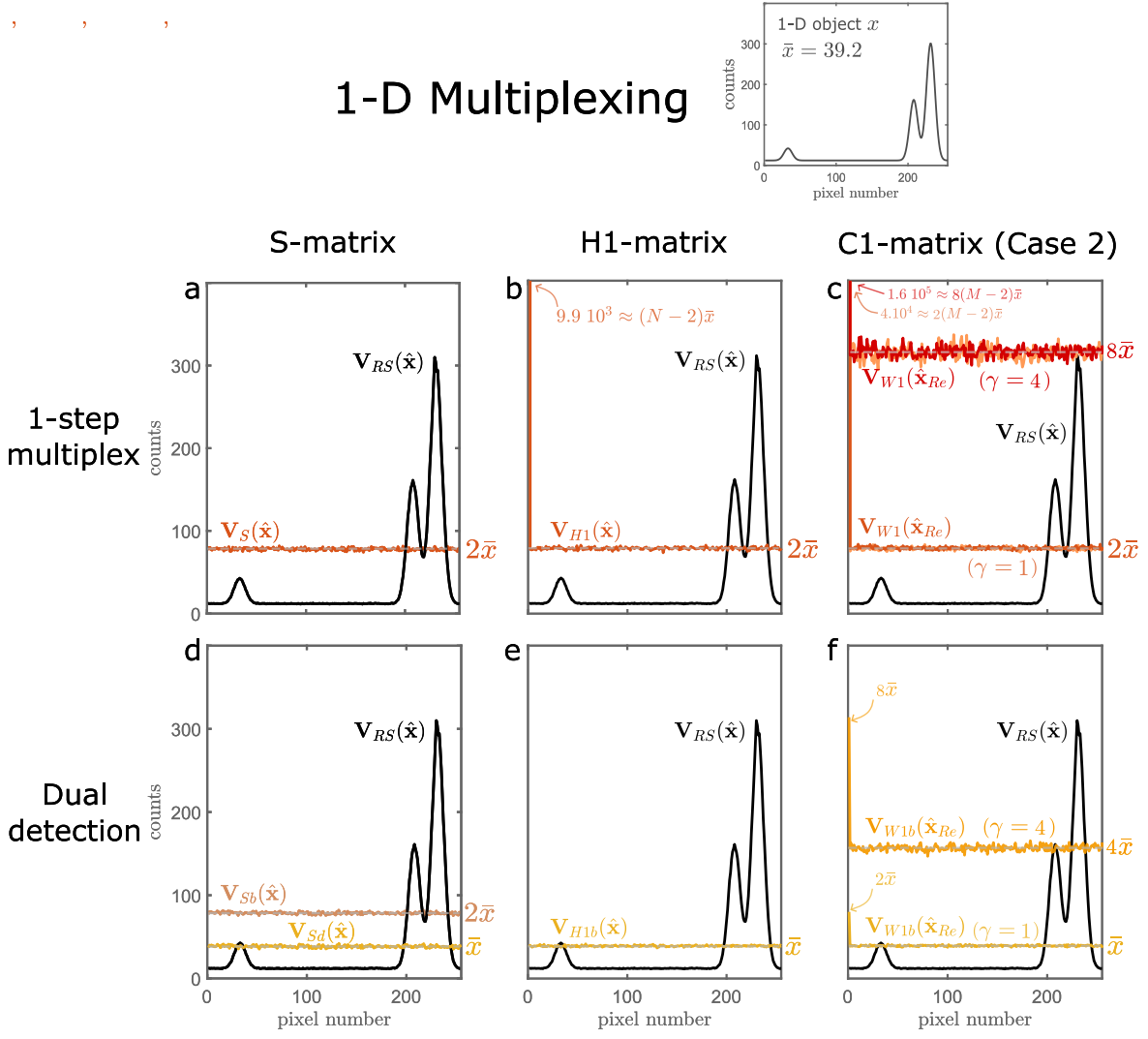


Figure 4.3: Empirical variances obtained in simulations with LS-estimation, for 1-step multiplexing and dual-detection, for the 3 considered multiplexing matrices, as compared to the variance associated with raster-scanning (black curves). For positive-cosine multiplexing, we show the resulting variance for  $\gamma = 1$  and  $\gamma = 4$ . (c) For positive-cosine multiplexing, the variance superimpose with the results obtained when multiplexing with a positive cosine matrix (light orange curves, see section 4.3.2). (d): For S-multiplexing, we find that the variance obtained with a balanced strategy ( $V_{Sb}(\hat{\mathbf{x}})$ ) is twice higher than when considering the full measurements ( $V_{Sd}(\hat{\mathbf{x}})$ ).  $N = 256$ ,  $M = 510$ , number of noise realisations = 5000.



## 2-D Multiplexing

2-D object  $x$   
 $\bar{x} = 43.5$

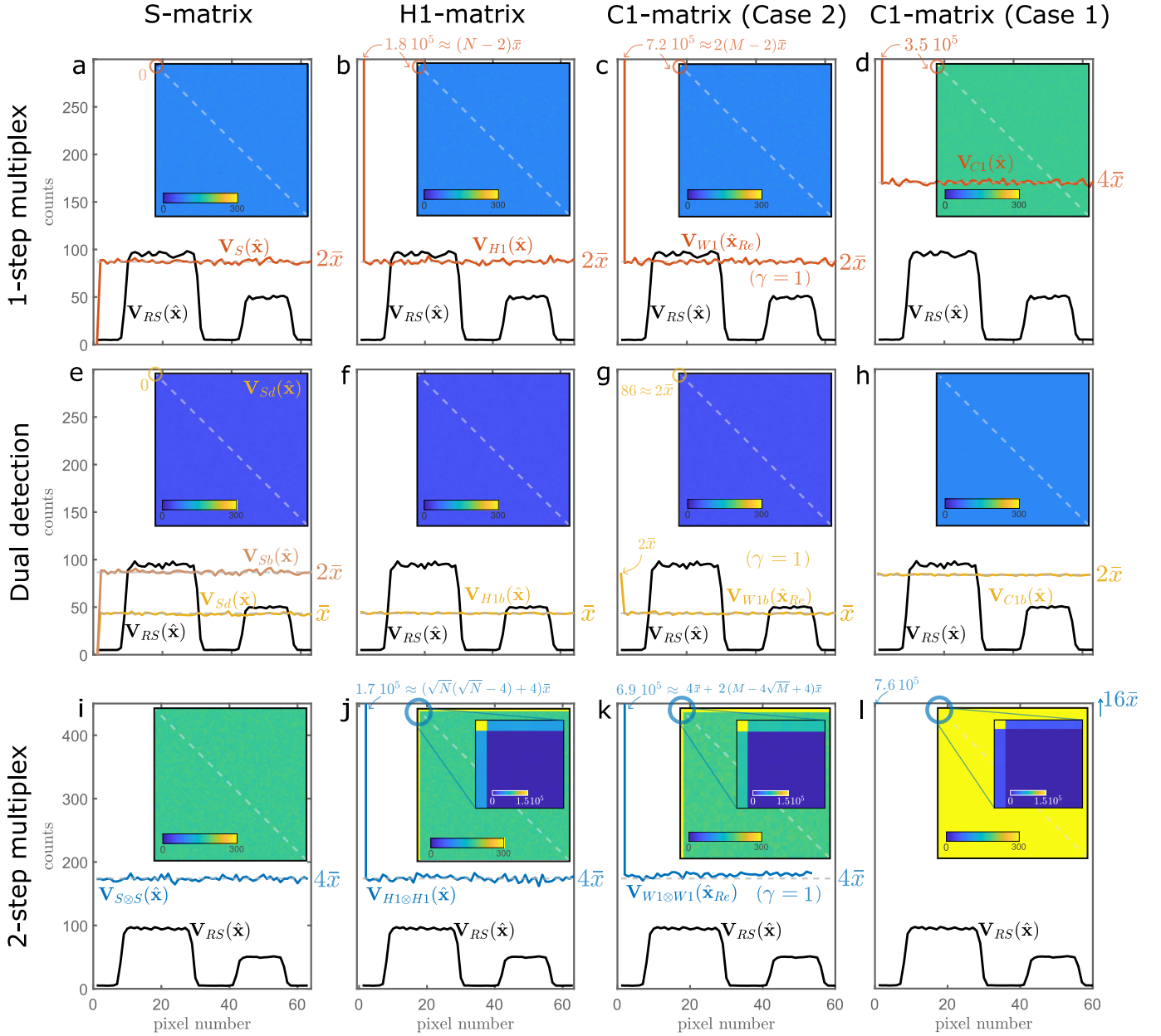


Figure 4.4: Empirical variances obtained in simulations with LS-estimation, for 1-step multiplexing, dual-detection and 2-step multiplexing, for several multiplexing matrices, as compared to the variance associated with raster-scanning (black curves). The images represent the 2-D variance along the pixels of the 2-D intensity object. The plots are the sections along the image diagonal. In (a) and (e), the first pixel variance is equal to zero, because we need to constrain the first object pixel to be zero to use the S-matrix of odd dimensions in 2-D (see paragraph below). In (e), the variance obtained with balanced detection is twice higher than when considering the full measurements. In (j-l), we zoom of the special pixels of the variance introduced by the structure of  $\mathbf{H}_1$  and  $\mathbf{C}_1$ .  $N = 4096$ ,  $M = 8190$ , number of noise realisations = 5000.

Some remarks on the implementation in 1-D and 2-D:

- The two-step multiplexing modality is not relevant in 1-D
- In 2-D, for one-step multiplexing, the 2-D patterns are obtained by reshaping each of rows of the multiplexing matrix. But in this work, the S-matrices are of odd dimensions (derived from Sylvester-Hadamard matrices). The S-matrix can be used in 2-D with a 'negative' Hadamard matrix  $\mathbf{H}_2$  (4.2.10)), if the first pixel of the object is zero and if the first measurement is discarded.

## Chapter 5

# Robustness to perturbations

In this chapter, we assess the robustness of the system to some perturbations that modify the initial equation (2.2.5):

$$\mathbf{b} \sim \text{Poisson}(\mathbf{A}\mathbf{x})$$

### 5.1 Initial model multiplied by a constant

Here, we consider a perturbation  $\epsilon > 0$  that modify the initial system such that:

$$\mathbf{b} \sim \text{Poisson}\left(\frac{1}{\epsilon}\mathbf{A}\mathbf{x}\right) \quad (5.1.1)$$

This accounts for various system losses or specific experimental details.

Often, this perturbation is not quantified and is simply incorporated in the object. Then, the quantity of interest is no longer the ground truth  $\mathbf{x}$  but a quantity proportional to it:  $\mathbf{y} = \frac{1}{\epsilon}\mathbf{x}$ . With this change of variable of interest, the initial model applies with  $\mathbf{b} \sim \text{Poisson}(\mathbf{A}\mathbf{y})$ ,  $\hat{\mathbf{y}} = \mathbf{A}^{-1}\mathbf{b}$ , and  $\mathbf{V}(\hat{\mathbf{y}}) = (\mathbf{A}^{-1} \odot \mathbf{A}^{-1})\mathbf{A}\mathbf{y} = \epsilon\mathbf{V}(\hat{\mathbf{x}})$ .

The perturbation can also be seen to impact the initial multiplexing matrix such that it is no longer  $\mathbf{A}$  but rather an equivalent matrix  $\tilde{\mathbf{A}}$ . In this case, equation (5.1.1) becomes:

$$\mathbf{b} \sim \text{Poisson}(\tilde{\mathbf{A}}\mathbf{x}) \quad (5.1.2)$$

with

$$\tilde{\mathbf{A}} = \frac{1}{\epsilon}\mathbf{A} \quad (5.1.3)$$

Using (2.2.9) directly leads to the associated estimation variance:

$$\tilde{\mathbf{V}}(\hat{\mathbf{x}}) = (\tilde{\mathbf{A}}^{-1} \odot \tilde{\mathbf{A}}^{-1})\tilde{\mathbf{A}}\mathbf{x} = \epsilon(\mathbf{A}^{-1} \odot \mathbf{A}^{-1})\mathbf{A}\mathbf{x} \quad (5.1.4)$$

i.e.

$$\boxed{\tilde{\mathbf{V}}(\hat{\mathbf{x}}) = \epsilon\mathbf{V}(\hat{\mathbf{x}})} \quad (5.1.5)$$

Where  $\mathbf{V}(\hat{\mathbf{x}})$  is the variance associated to  $\mathbf{A}$  given in equation (2.2.9).

#### 5.1.1 Implications for the three multiplexing schemes

The implications for of this result for the three positive-multiplexing schemes are listed in the table below.

|  | One-step multi-plexing                 | Two-step multiplexing                  |   | Dual-detection  |
|--|--|--|---|---|
| $\tilde{\mathbf{A}}$                   | $\frac{1}{\epsilon}\mathbf{A}$         | $\frac{1}{\epsilon}\mathbf{A}$         | $\frac{1}{\epsilon_p}\mathbf{P}^T \otimes \frac{1}{\epsilon_u}\mathbf{U}$ | $\frac{1}{\epsilon} \left[ \mathbf{A}_1 \ \mathbf{A}_2 \right]^T$ |
| $\tilde{\mathbf{V}}(\hat{\mathbf{x}})$ | $\epsilon\mathbf{V}(\hat{\mathbf{x}})$ | $\epsilon\mathbf{V}(\hat{\mathbf{x}})$ | $\epsilon_p\epsilon_u\mathbf{V}(\hat{\mathbf{x}})$                        | $\epsilon\mathbf{V}(\hat{\mathbf{x}})$                            |

Table 5.1: Variances for different perturbations of the multiplexing matrix.

### Implications for one-step multiplexing

In a one-step multiplexing scenario, the above equation (5.1.5) directly applies.

### Implications for two-step multiplexing

In a two-step multiplexing scheme (section 3.2), the perturbation may happen on the global multiplexing matrix  $\mathbf{A} = \mathbf{P}^T \otimes \mathbf{U}$ , in which case the above result (5.1.5) applies. On opposite, if two perturbations are applied on each sub-matrix such that:  $\tilde{\mathbf{P}} = \frac{1}{\epsilon_p}\mathbf{P}$  and  $\tilde{\mathbf{U}} = \frac{1}{\epsilon_u}\mathbf{U}$ , the equivalent multiplexing matrix is:

$$\tilde{\mathbf{A}} = \frac{1}{\epsilon_p\epsilon_u}\mathbf{A} \quad (5.1.6)$$

Hence the resulting variance reads:

$$\tilde{\mathbf{V}}(\hat{\mathbf{x}}) = \epsilon_p\epsilon_u\mathbf{V}(\hat{\mathbf{x}}) \quad (5.1.7)$$

Where  $\mathbf{V}(\hat{\mathbf{x}})$  is the variance associated to  $\mathbf{A}$  given in equation (3.2.5).

### Implications for dual-detection

In the dual detection modality, the perturbation can modelled via the following equivalent matrix:

$$\tilde{\mathbf{A}} = \frac{1}{\epsilon} \left[ \mathbf{A}_1 \ \mathbf{A}_2 \right]^T = \frac{1}{\epsilon}\mathbf{A} \quad (5.1.8)$$

In this case we also have:

$$\tilde{\mathbf{V}}(\hat{\mathbf{x}}) = \epsilon\mathbf{V}(\hat{\mathbf{x}}) \quad (5.1.9)$$

Where  $\mathbf{V}(\hat{\mathbf{x}})$  is the variance associated to  $\mathbf{A}$  given in equation (3.3.10).

Proof: Adapting equation (3.3.10) yields:

$$\mathbf{V}(\hat{\mathbf{x}}) = (\tilde{\mathbf{A}}^+ \odot \tilde{\mathbf{A}}^+)\tilde{\mathbf{A}}\mathbf{x} \quad (5.1.10)$$

with

$$\tilde{\mathbf{A}}^+ = (\tilde{\mathbf{A}}^T\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^T = \left( \frac{1}{\epsilon^2}\mathbf{A}^T\mathbf{A} \right)^{-1} \times \frac{1}{\epsilon}\mathbf{A}^T = \gamma(\mathbf{A}^T\mathbf{A})\mathbf{A}^T = \epsilon\mathbf{A}^+ \quad (5.1.11)$$

Combining the two above equations leads to the results of equation (5.1.9).

### 5.1.2 Implications for comparisons at constant number of photons

In this work, we focus on comparing positive-multiplexing and raster-scanning when the number of photons between is *not* constant. Then, for  $N$  measurements, raster-scanning leads to a total of  $N\bar{x}$  photon counts, and positive-multiplexing to a total of about  $\frac{N^2}{a}\bar{x}$  photon counts (where  $a$  depends on the multiplexing matrix).

Yet, the above result of equation (5.1.5) also allows us to compare the variance of raster-scanning and positive-multiplexing at constant number of photons (same number of photons detected in the

two cases). Comparing raster-scanning and positive-multiplexing at constant number of photons is equivalent to divide the number of collected photons by multiplexing  $\frac{N}{a}$  and thus multiplex with an equivalent matrix  $\tilde{\mathbf{A}} = \frac{a}{N}\mathbf{A}$ . Then, adapting equation 5.1.5 results in a positive-multiplexing variance of:

$$\tilde{\mathbf{V}}_{mult}(\hat{\mathbf{x}}) = \frac{N}{a}\mathbf{V}(\hat{\mathbf{x}}) \quad (5.1.12)$$

where  $\mathbf{V}(\hat{\mathbf{x}})$  is the variance when the number of photons is  $\frac{N^2}{a}\bar{x}$ . The variance associated with raster-scanning stays  $\mathbf{V}_{rs}(\hat{\mathbf{x}}) = \mathbf{x}$ .

Therefore, performing positive-multiplexing with fewer photons worsen the associated variance by the same factor: the estimation variance is worsened by a factor proportional to  $N$ , and thus the SNR by a factor proportional  $\sqrt{N}$ .

Note that in positive-Hadamard-based and positive-Cosine multiplexing,  $a = 2$  and  $\mathbf{V}(\hat{\mathbf{x}}) = k\bar{x}$ , therefore  $\tilde{\mathbf{V}}_{mult}(\hat{\mathbf{x}}) = \frac{N}{2}k\bar{x}$ .

## 5.2 Initial model with additional noise or background

In this section, we consider some additional noise sources or unwanted background signal that may arise in the experimental system (Fig. 5.1):

- Additional white Gaussian noise (AWGN)  $\mathbf{e}$  at the detection (e.g. electronic noise arising from the detector)
- Additional known background  $\boldsymbol{\eta}$  that does not experience the multiplexing step
- Additional known background  $\boldsymbol{\beta}$  that does experience the multiplexing step
- Additional constant offset  $\alpha$  in the multiplexing matrix

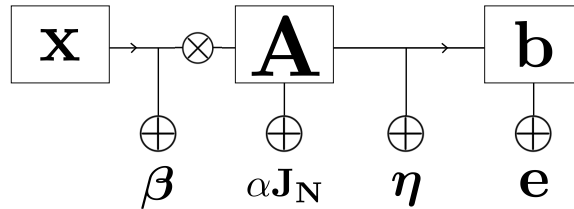


Figure 5.1: Schematic representation of some sources of noise or unwanted background that may arise during a measurement.  $\mathbf{x}$ : object,  $\mathbf{A}$ : multiplexing matrix,  $\mathbf{b}$ : measurement,  $\boldsymbol{\beta}$ : unwanted background or signal added to the system before the multiplexing stage,  $\alpha\mathbf{J}_N$ : unwanted constant offset added to the multiplexing matrix,  $\boldsymbol{\eta}$ : unwanted background or signal added to the system after the multiplexing stage,  $\mathbf{e}$ : detector electronic noise

For each scenario, we derive the variances associated to raster-scanning, and to positive-multiplexing with a general matrix  $\mathbf{A}$ . We also give the results for a special class of matrices  $\mathbf{A}_c$  that lead to a constant variance:

$$\mathbf{V}_{Ac}(\hat{\mathbf{x}}) = k\bar{x}\mathbf{1}_N$$

Details on this class of matrices are provided in chapter 6.  $\mathbf{A}_c$  is defined such that  $\mathbf{A}_c \in \mathbb{R}^{N^2}$  (for  $N > 2$ ) with an inverse of:

$$\mathbf{A}_c^{-1} = \frac{k}{N}(N - 2p)\mathbf{\Lambda}$$

where  $\mathbf{\Lambda}$  is a matrix with elements  $\delta_{ij} = \pm 1$ , in which each column contains  $p$  negative elements (equation (6.2.6)).

### 5.2.1 Results

The obtained variance for each scenario are summarized in the table below. The proofs are provided in section 5.2.2. The results of the table are illustrated in Fig. 5.2 for S-multiplexing ( $\mathbf{A}_c = \mathbf{S}$ -matrix, i.e.  $k = \frac{2N}{N+1} \approx 2$  and  $q = 1$ ).

|  | AWGN $\mathbf{e}$   | Known bck $\boldsymbol{\eta}$   | Known bck $\boldsymbol{\beta}$   | Matrix with constant offset $\alpha$   |
|--|---|---|--|--|
| $\mathbf{b}_0$   | $\mathbf{A}\mathbf{x}$  | $\mathbf{A}\mathbf{x} + \boldsymbol{\eta}$  | $\mathbf{A}(\mathbf{x} + \boldsymbol{\beta})$  | $(\mathbf{A} + \alpha\mathbf{J}_N)\mathbf{x}$  |
| $\mathbf{b}$   | $Poisson(\mathbf{A}\mathbf{x}) + \mathbf{e}$  | $Poisson(\mathbf{A}\mathbf{x} + \boldsymbol{\eta})$                                       | $Poisson(\mathbf{A}(\mathbf{x} + \boldsymbol{\beta}))$   | $Poisson((\mathbf{A} + \alpha\mathbf{J}_N)\mathbf{x})$   |
| $\hat{\mathbf{x}}_{LS}$  | $\mathbf{A}^{-1}\mathbf{b}$   | $\mathbf{A}^{-1}(\mathbf{b} - \boldsymbol{\eta})$   | $\mathbf{A}^{-1}\mathbf{b} - \boldsymbol{\beta}$   | $(\mathbf{A} + \alpha\mathbf{J}_N)^{-1}\mathbf{b}$   |
| General case   |   |   |  |  |
| $\tilde{\mathbf{V}}(\hat{\mathbf{x}})$   | $\mathbf{V}(\hat{\mathbf{x}}) + \sigma^2 \text{diag}((\mathbf{A}^T \mathbf{A})^{-1})$ | $\mathbf{V}(\hat{\mathbf{x}}) + (\mathbf{A}^{-1} \odot \mathbf{A}^{-1})\boldsymbol{\eta}$ | $\mathbf{V}(\hat{\mathbf{x}}) + (\mathbf{A}^{-1} \odot \mathbf{A}^{-1})\mathbf{A}\boldsymbol{\beta}$ | $((\mathbf{A} + \alpha\mathbf{J}_N)^{-1}) \odot (\mathbf{A} + \alpha\mathbf{J}_N)^{-1}(\mathbf{A} + \alpha\mathbf{J}_N)\mathbf{x}$ |
| Raster-scanning  |   |   |  |  |
| $\tilde{V}_{RS}(\hat{x}_i)$  | $x_i + \sigma^2$  | $x_i + \eta_i$  | $x_i + \beta_i$  | $x_i + \alpha N \bar{x}$ (if $\alpha \ll 1$ )  |
| Positive-multiplexing with $\mathbf{A} = \mathbf{A}_c$ (such that $V_{Ac}(\hat{x}_i) = k\bar{x}$ ) |   |   |  |  |
| $\tilde{V}_{Ac}(\hat{x}_i)$  | $k\bar{x} + \frac{k^2 q^2}{N} \sigma^2$   | $k\bar{x} + \frac{k^2 q^2}{N} \bar{\eta}$   | $k\bar{x} + k\bar{\beta}$  | $k\bar{x} + \alpha k^2 q^2 \bar{x}$<br>(if $N \gg 1$ )   |

Table 5.2: Resulting variances in the presence of noise or unwanted signal. AWGN: Additive White Gaussian Noise;  $\sigma$ : standard deviation of  $\mathbf{e}$ ;  $\mathbf{b}_0$ : general noiseless model,  $\mathbf{b}$ : noise model,  $\hat{\mathbf{x}}_{LS}$ : least-square estimate,  $\mathbf{V}(\hat{\mathbf{x}})$ : associated variance.  $k$ : real positive constant.  $q = (N - 2p)$ : defined in equation (6.2.6).  $\eta_i \geq 0$ ;  $\beta_i \geq 0$  and  $\alpha$  are supposed to be known from a calibration step. The proofs and exact expressions are provided in the text.

Overall, raster-scanned measurements are highly sensitive to the four considered additional perturbations: the additional noise variance adds to the signal. In opposite, in the two first scenarios, positive-multiplexing with  $\mathbf{A}_c$  is very robust to the perturbations  $\mathbf{e}$  and  $\boldsymbol{\eta}$ : they hardly impact the variance if  $N \gg 1$ . The stronger the perturbation, the more object pixels are better estimated with  $\mathbf{A}_c$ -multiplexing (as compared to raster-scanning). Note that in the limit case where  $\sigma^2 \geq k\bar{x}$  or  $\eta_i \geq k\bar{x}$ ,  $\mathbf{A}_c$ -multiplexing brings a SNR advantage over raster-scanning on all pixels. The effect is similar in the last scenario. The results are different in the third scenario, i.e. if some background signal  $\boldsymbol{\beta}$  is added to the system before the multiplexing step. There, the perturbation impacts on average  $k$  times more  $\mathbf{A}_c$ -multiplexing than raster-scanning. Then, the stronger the perturbation, the more object pixels are better estimated with raster-scanning (as compared to  $\mathbf{A}_c$ -multiplexing).

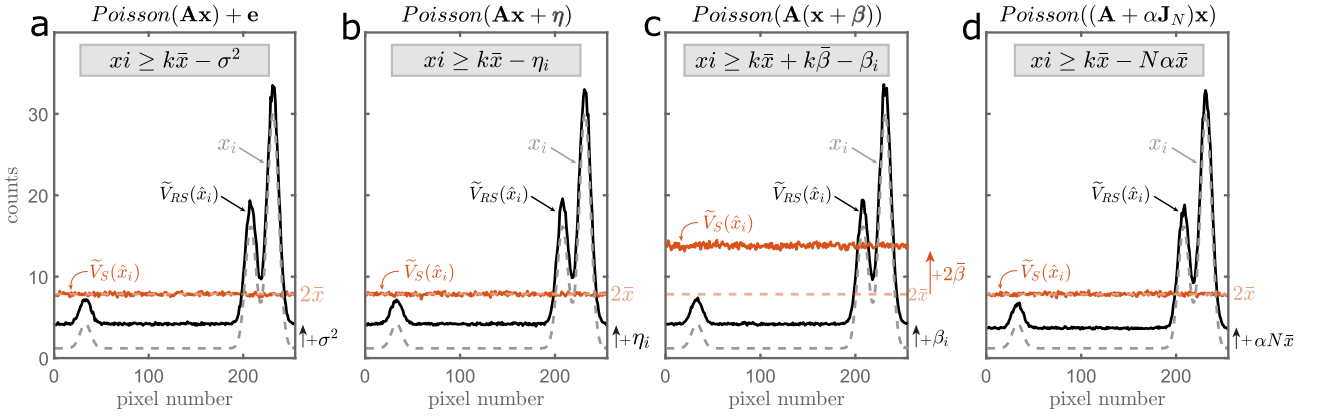


Figure 5.2: Empirical variances obtained for raster-scanning (black) and positive multiplexing with  $\mathbf{A}_c = \mathbf{S}()$ , for the four above scenarios and a 1-D object. For comparison, the theoretical variances that would be obtained in the initial shot-noise limited model (equation (2.2.5)) are indicated in grey and light red dashed lines for raster-scanning and S-multiplexing respectively. The gray boxes indicate the conditions on object pixel  $i$  for Ac-multiplexing to bring an SNR advantage over raster-scanning (valid for for  $N \gg 1$ ,  $\alpha \ll 1$  and  $q$  independent of  $N$ ). In the simulation,  $\eta$  and  $\beta$  are constant backgrounds,  $\sigma^2 = \eta_i = \beta_i = 3$  counts;  $N\alpha\bar{x} = 2.6$  counts; with  $\alpha = 0.0025$ ,  $\bar{x} = 4$  counts;  $N = 255$  pixels. Number of noise realisations = 5000.

## 5.2.2 Details and proofs

### Additional electronic noise $\mathbf{e}$

With additive electronic noise, the model becomes:

$$\mathbf{b} \sim \text{Poisson}(\mathbf{A}\mathbf{x}) + \mathbf{e} \quad (5.2.13)$$

The electronic noise is modelled as additive white gaussian noise (AWGN)  $\mathbf{e}$ , with  $\forall i \neq j, E(e_i) = 0$ ,  $E(e_i e_j) = 0$ ,  $E(e_i^2) = \sigma^2$ , where  $\sigma$  is the noise standard deviation.

To simplify the calculation, we re-write  $\mathbf{b}$  as

$$\mathbf{b} = \mathbf{p} + \mathbf{e} \quad (5.2.14)$$

where  $\mathbf{p}$  approximates the Poisson distribution.  $\mathbf{p}$  follows a Gaussian distribution with mean equals its variance  $E(p_i) = E(p_i^2) = (\mathbf{A}\mathbf{x})_i$ . The error  $\delta\hat{\mathbf{x}}$  reads:

$$\delta\hat{\mathbf{x}} = \mathbf{A}^{-1}(\mathbf{p} + \mathbf{e}) \quad (5.2.15)$$

This leads to the covariance matrix:

$$\begin{aligned} \mathbf{\Gamma} &= \mathbf{A}^{-1}(\langle \mathbf{p}\mathbf{p}^T \rangle + \langle \mathbf{e}\mathbf{e}^T \rangle) \mathbf{A}^{-T} \text{ (uncorrelated noise)} \\ &= \mathbf{A}^{-1}(\text{Diag}(\mathbf{A}\mathbf{x}) + \sigma^2 \mathbf{I}_N) \mathbf{A}^{-T} \end{aligned}$$

The resulting variance is then:

$$\tilde{\mathbf{V}}(\hat{\mathbf{x}}) = (\mathbf{A}^{-1} \odot \mathbf{A}^{-1}) \mathbf{A}\mathbf{x} + \sigma^2 \text{diag}((\mathbf{A}^T \mathbf{A})^{-1}) \quad (5.2.16)$$

Hence, for raster-scanning, the variance reads:

$$\tilde{\mathbf{V}}_{RS}(\hat{\mathbf{x}}) = \mathbf{x} + \sigma^2 \mathbf{1}_N \quad (5.2.17)$$

and for multiplexing with  $\mathbf{A}_c$ , the variance reads:

$$\tilde{\mathbf{V}}_{Ac}(\hat{\mathbf{x}}) = k\bar{x}\mathbf{1}_N + \sigma^2 \text{diag}((\mathbf{A}_c^T \mathbf{A}_c)^{-1})$$

$$\boxed{\tilde{\mathbf{V}}_{Ac}(\hat{\mathbf{x}}) = (k\bar{x} + \frac{k^2q^2}{N}\sigma^2)\mathbf{1}_N} \quad (5.2.18)$$

with  $q = (N - 2p)$ . If  $q$  is independent of  $N$  (true for matrices  $\mathbf{S}$ ,  $\mathbf{H}_1$  and  $\mathbf{C}_1$ ) and  $N \gg 1$ :

$$\tilde{\mathbf{V}}_{Ac}(\hat{\mathbf{x}}) \approx k\bar{x}\mathbf{1}_N$$

Then, under these assumptions positive-multiplexing with  $\mathbf{A}_c$  is advantageous over raster-scanning on pixels  $i$  for which:

$$\boxed{x_i \geq k\bar{x} - \sigma^2}$$

Example for the  $\mathbf{S}$ -matrix:  $k = 2N/(N + 1)$  and  $q = 1$ , thus:

$$\mathbf{V}_S(\hat{\mathbf{x}}) = \frac{2N}{N+1}\bar{x} + \frac{4N}{(N+1)^2}\mathbf{1}_N \approx (2\bar{x} + \frac{4\sigma^2}{N})\mathbf{1}_N \approx 2\bar{x}\mathbf{1}_N \quad (\text{for } N \gg 1)$$

### Additional non-multiplexed known background $\eta$

This case resembles the above case. The background does not experiences the multiplexing matrix. It can depend on the object or not, and is supposed to be known from a calibration step. The model is:

$$\mathbf{b} \sim \text{Poisson}(\mathbf{A}\mathbf{x} + \eta) \quad (5.2.19)$$

The error  $\delta\hat{\mathbf{x}}$  for the LS estimate reads:

$$\delta\hat{\mathbf{x}} = \mathbf{A}^{-1}(\mathbf{b} - \eta) - \mathbf{A}^{-1}(\mathbf{b}_0 - \eta) = \mathbf{A}^{-1}\delta\mathbf{b} \quad (5.2.20)$$

and leads to the covariance matrix:

$$\mathbf{\Gamma} = \mathbf{A}^{-1}\text{Diag}(\mathbf{A}\mathbf{x} + \eta)\mathbf{A}^{-T} \quad (5.2.21)$$

Thus:

$$\boxed{\tilde{\mathbf{V}}(\hat{\mathbf{x}}) = (\mathbf{A}^{-1} \odot \mathbf{A}^{-1})\mathbf{A}\mathbf{x} + (\mathbf{A}^{-1} \odot \mathbf{A}^{-1})\eta} \quad (5.2.22)$$

For raster-scanning,

$$\boxed{\tilde{\mathbf{V}}_{RS}(\hat{\mathbf{x}}) = \mathbf{x} + \eta} \quad (5.2.23)$$

For multiplexing with  $\mathbf{A}_c$ , the variance reads:

$$\tilde{\mathbf{V}}_{Ac}(\hat{\mathbf{x}}) = k\bar{x}\mathbf{1}_N + \frac{k^2}{N^2}(N - 2p)^2\mathbf{J}_N\eta$$

$$\boxed{\tilde{\mathbf{V}}_{Ac}(\hat{\mathbf{x}}) = (k\bar{x} + \frac{k^2q^2}{N}\bar{\eta})\mathbf{1}_N} \quad (5.2.24)$$

with  $q = (N - 2p)$ . If  $q^2$  is independent of  $N$  (true for matrices  $\mathbf{S}$ ,  $\mathbf{H}_1$  and  $\mathbf{C}_1$ ) and  $N \gg 1$ :

$$\tilde{\mathbf{V}}_{Ac}(\hat{\mathbf{x}}) \approx k\bar{x}\mathbf{1}_N \quad (5.2.25)$$

In this case, multiplexing with  $\mathbf{A}_c$  is advantageous over raster-scanning for pixels  $i$  for which:

$$\boxed{x_i \geq k\bar{x} - \eta_i} \quad (5.2.26)$$

Example for the  $\mathbf{S}$ -matrix:  $k = 2N/(N + 1)$  and  $q = 1$ , thus:

$$\tilde{\mathbf{V}}_S(\hat{\mathbf{x}}) = \frac{2N}{N+1}\bar{x}\mathbf{1}_N + \frac{4N}{(N+1)^2}\bar{\eta}\mathbf{1}_N \approx (2\bar{x} + \frac{4}{N}\bar{\eta})\mathbf{1}_N \approx 2\bar{x}\mathbf{1}_N \quad (\text{for } N \gg 1)$$



### Additional multiplexed known background $\beta$

Here, there is again some known background that may arise from different experimental sources. It experiences the multiplexing matrix, which means it is added to the system before the multiplexing step. It may again depend on the object or not. The model reads:

$$\mathbf{b} \sim \text{Poisson}(\mathbf{A}(\mathbf{x} + \beta)) \quad (5.2.27)$$

The error  $\delta\hat{\mathbf{x}}$  reads:

$$\delta\hat{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{b} - \beta - (\mathbf{A}^{-1}\mathbf{b}_0 - \beta) = \mathbf{A}^{-1}\delta\mathbf{b} \quad (5.2.28)$$

This leads to the covariance matrix:

$$\mathbf{\Gamma} = \mathbf{A}^{-1} \text{Diag}(\mathbf{A}(\mathbf{x} + \beta)) \mathbf{A}^{-T} \quad (5.2.29)$$

and to the variance:

$$\boxed{\tilde{\mathbf{V}}(\hat{\mathbf{x}}) = (\mathbf{A}^{-1} \odot \mathbf{A}^{-1}) \mathbf{A} \mathbf{x} + (\mathbf{A}^{-1} \odot \mathbf{A}^{-1}) \mathbf{A} \beta} \quad (5.2.30)$$

For raster-scanning:

$$\boxed{\tilde{\mathbf{V}}_{RS}(\hat{\mathbf{x}}) = \mathbf{x} + \beta} \quad (5.2.31)$$

For multiplexing with  $\mathbf{A}_c$ , the variance reads:

$$\boxed{\tilde{\mathbf{V}}_{Ac}(\hat{\mathbf{x}}) = (k\bar{x} + k\bar{\beta}) \mathbf{1}_N} \quad (5.2.32)$$

Therefore, here multiplexing with  $\mathbf{A}_c$  is advantageous over raster-scanning for pixels  $i$  for which:

$$\boxed{x_i \geq k\bar{x} + k\bar{\beta} - \beta_i} \quad (5.2.33)$$

Or, if  $\beta_i = \bar{\beta} \forall i$ :

$$x_i \geq k\bar{x} + (k-1)\bar{\beta} \quad (5.2.34)$$

Example for the  $\mathbf{S}$ -matrix:  $k = 2N/(N+1)$ . If  $\beta_i = \bar{\beta} \forall i$ , then:

$$\mathbf{V}_S(\hat{\mathbf{x}}) = \frac{2N}{N+1} \bar{x} \mathbf{1}_N + \frac{2N}{N+1} \bar{\beta} \mathbf{1}_N \approx (2\bar{x} + 2\bar{\beta}) \mathbf{1}_N \quad \text{for } N \gg 1$$

### Additional constant offset $\alpha$ on the multiplexing matrix

This scenario represents the case where, instead of multiplexing by the planned matrix  $\mathbf{A}$ , the object is actually multiplexed by this matrix plus an offset  $\alpha$  that depends on the object signal. It is the proportion of the total signal from the object that contributes to the measurement when it is not expected to be. The model is:

$$\mathbf{b} \sim \text{Poisson}((\mathbf{A} + \alpha \mathbf{J}_N) \mathbf{x}) = \text{Poisson}(\mathbf{C} \mathbf{x}) \quad (5.2.35)$$

with  $\mathbf{C} = \mathbf{A} + \alpha \mathbf{J}_N$ . The variance thus derives from the initial model formula of equation (2.2.9):

$$\boxed{\tilde{\mathbf{V}} = ((\mathbf{A} + \alpha \mathbf{J}_N)^{-1}) \odot ((\mathbf{A} + \alpha \mathbf{J}_N)^{-1}) ((\mathbf{A} + \alpha \mathbf{J}_N) \mathbf{x})} \quad (5.2.36)$$

For raster-scanning,  $\mathbf{C}_{rs} = \mathbf{I}_N + \alpha \mathbf{J}_N$  and  $\mathbf{C}_{RS}^{-1} = \mathbf{I}_N - \frac{\alpha \mathbf{J}_N}{1 + \alpha N}$  (equation (0.0.3)). Thus,

$$\begin{aligned} \tilde{\mathbf{V}}_{RS}(\hat{\mathbf{x}}) &= (\mathbf{C}_{RS}^{-1} \odot \mathbf{C}_{RS}^{-1}) (\mathbf{C}_{RS} \mathbf{x}) \\ &= (\mathbf{I}_N - \frac{\alpha \mathbf{J}_N}{1 + \alpha N})^{\odot 2} (\mathbf{I}_N + \alpha \mathbf{J}_N) \mathbf{x} \\ &= (\mathbf{I}_N - \frac{2\alpha}{1 + \alpha N} \mathbf{I}_N + \frac{\alpha^2}{(1 + \alpha N)^2} \mathbf{J}_N) (\mathbf{I}_N + \alpha \mathbf{J}_N) \mathbf{x} \\ &= (1 - \frac{2\alpha}{1 + \alpha N}) \mathbf{x} + (\alpha - \frac{2\alpha^2}{1 + \alpha N} + \frac{\alpha^2 + \alpha^3 N}{(1 + \alpha N)^2}) N \bar{x} \mathbf{1}_N \\ &= (1 - \frac{2\alpha}{1 + \alpha N}) \mathbf{x} + \frac{\alpha N ((1 + \alpha N)^2 - \alpha(1 + \alpha N))}{(1 + \alpha N)^2} \bar{x} \\ &= (1 - \frac{2\alpha}{1 + \alpha N}) \mathbf{x} + \frac{\alpha N (1 + \alpha N - \alpha)}{1 + \alpha N} \bar{x} \end{aligned}$$

and, if  $\alpha \ll 1$  or  $N \gg 1$ :

$$\boxed{\tilde{\mathbf{V}}_{RS}(\hat{\mathbf{x}}) \approx \mathbf{x} + \alpha N \bar{x} \mathbf{1}_N} \quad (5.2.37)$$

For multiplexing with  $\mathbf{A}_c$ ,  $\mathbf{C}_{Ac} = \mathbf{A}_c + \alpha \mathbf{J}_N$ . We use the Sherman-Morrison formula (equation (0.0.3)) to calculate the inverse of  $\mathbf{C}_A$ :

$$\begin{aligned} \mathbf{C}_{Ac}^{-1} &= \mathbf{A}_c^{-1} - \frac{\alpha \mathbf{A}_c^{-1} \mathbf{J}_N \mathbf{A}_c^{-1}}{1 + \alpha \mathbf{1}^T \mathbf{A}_c^{-1} \mathbf{1}} \\ &= \mathbf{A}_c^{-1} - \frac{\alpha k (N - 2p)^2}{N(1 + \alpha k (N - 2p)^2)} \mathbf{A}_c^{-1} \mathbf{J}_N \\ &= \mathbf{A}_c^{-1} \left( \mathbf{I}_N - \frac{\alpha k q^2}{N(1 + \alpha k q^2)} \mathbf{J}_N \right) \end{aligned}$$

with  $q = (N - 2p)$ . We would need a further condition on the sum of the rows of  $\mathbf{A}$  to be able to carry out the derivation. If  $N \gg 1$  and  $q^2$  is independent of  $N$  (true for matrices  $\mathbf{S}$ ,  $\mathbf{H}_1$  and  $\mathbf{C}_1$ ), then  $\mathbf{C}_{Ac}^{-1} \approx \mathbf{A}_c^{-1}$  and:

$$\begin{aligned} \tilde{\mathbf{V}}_{Ac}(\hat{\mathbf{x}}) &\approx (\mathbf{A}_c^{-1} \odot \mathbf{A}_c^{-1})(\mathbf{A}_c + \alpha \mathbf{J}_N) \mathbf{x} \\ &= k \bar{x} \mathbf{1}_N + \alpha \frac{k}{N} \mathbf{J}_N \mathbf{A}_c^{-1} \mathbf{J}_N \mathbf{x} \\ &= k \bar{x} \mathbf{1}_N + \alpha \frac{k^2 q^2}{N^2} \mathbf{J}_N \mathbf{J}_N \mathbf{x} \end{aligned}$$

i.e.

$$\boxed{\tilde{\mathbf{V}}_{Ac}(\hat{\mathbf{x}}) \approx (k + \alpha k^2 q^2) \bar{x} \mathbf{1}_N} \quad (5.2.38)$$

if in addition,  $\alpha \ll 1$ , multiplexing with  $\mathbf{A}_c$  is advantageous over raster-scanning for pixels  $i$  for which:

$$\boxed{x_i \geq k \bar{x} - N \alpha \bar{x}} \quad (5.2.39)$$

Example of the  $\mathbf{S}$ - matrix:  $\mathbf{C}_S = \mathbf{S} + \alpha \mathbf{J}_N$  and its inverse reads:

$$\mathbf{C}_S^{-1} = \mathbf{S}^{-1} - \frac{\alpha \mathbf{S}^{-1} \mathbf{J}_N \mathbf{S}^{-1}}{1 + \alpha \mathbf{1}^T \mathbf{S}^{-1} \mathbf{1}} = \mathbf{S}^{-1} - \frac{4\alpha}{(N+1)(1+N+2\alpha N)} \mathbf{J}_N \quad (5.2.40)$$

where we used the fact that  $\mathbf{J}_N \mathbf{S}^{-1} = \mathbf{S}^{-1} \mathbf{J}_N = 2/(N+1)$ , and that the sum of all elements of  $\mathbf{S}^{-1}$ , i.e.  $\mathbf{1}^T \mathbf{S}^{-1} \mathbf{1}$  is  $2N/(N+1)$ . Then, using equations (4.2.7) and (4.2.5) leads to:

$$\begin{aligned} \tilde{\mathbf{V}}_S(\hat{\mathbf{x}}) &= (\mathbf{C}_S^{-1} \odot \mathbf{C}_S^{-1})(\mathbf{C}_S \mathbf{x}) \\ &= \left( \mathbf{S}^{-1} - \frac{4\alpha}{(N+1)(1+N+2\alpha N)} \mathbf{J}_N \right)^{\odot 2} (\mathbf{S}^{-1} + \alpha \mathbf{J}_N) \mathbf{x} \\ &= \left( \frac{4}{(N+1)^2} \left( 1 + \frac{4\alpha^2}{(1+N+2\alpha N)^2} \right) \mathbf{J}_N - \frac{8\alpha}{(N+1)(1+N+2\alpha N)} \right) \mathbf{S}^{-1} (\mathbf{S} + \alpha \mathbf{J}_N) \mathbf{x} \\ &= \left( \left( \frac{2}{N+1} + \frac{4\alpha N}{(N+1)^2} \right) \left( 1 + \frac{4\alpha^2}{(1+N+2\alpha N)^2} \right) - \frac{16\alpha^2}{(N+1)^2(1+N+2\alpha N)} \right) \mathbf{J}_N \mathbf{x} \\ &\quad - \frac{8\alpha}{(N+1)(1+N+2\alpha N)} \mathbf{I}_N \mathbf{x} \end{aligned}$$

Then, for  $N \gg 1$ :

$$\begin{aligned} \tilde{\mathbf{V}}_S(\hat{\mathbf{x}}) &\approx \left( \left( \frac{2}{N} + \frac{4\alpha N}{N^2} \right) \left( 1 + \frac{4\alpha^2}{(N+2\alpha N)^2} \right) - \frac{16\alpha^2}{N^2(N+2\alpha N)} \right) N \bar{x} \mathbf{1}_N - \frac{8\alpha}{N(N+2\alpha N)} \mathbf{x} \\ &\approx (2(1+2\alpha) \left( 1 + \frac{4\alpha^2}{N^2(1+2\alpha)^2} \right) - \frac{16\alpha^2}{N^2(1+2\alpha)}) \bar{x} \mathbf{1}_N - \frac{8\alpha}{N^2(1+2\alpha)} \mathbf{x} \\ &\approx (2+4\alpha) \bar{x} \mathbf{1}_N - \frac{8\alpha}{N^2(1+2\alpha)} (\mathbf{x} + \bar{x} \mathbf{1}_N) \end{aligned}$$

Therefore, if  $N \gg 1$ :

$$\tilde{\mathbf{V}}_S(\hat{\mathbf{x}}) \approx (2 + 4\alpha)\bar{x}\mathbf{1}_N \quad (5.2.41)$$

In this case, S-multiplexing is therefore advantageous for pixels  $i$  for which:

$$\begin{aligned} x_i + \alpha N \bar{x} &\geq 2\bar{x} + 4\alpha\bar{x} \quad \text{i.e.} \\ x_i &\geq 2\bar{x} - N\alpha\bar{x} \end{aligned}$$

## Chapter 6

# Some facts on matrices leading to a constant variance

In chapter 4, we found that for three common multiplexing matrices, namely the **S**-matrix, the **H1**-matrix and the **W1**-matrix, when  $N \gg 1$ , the estimation variance reads:  $V(\hat{x}_i) \approx k\bar{x}$  on most object pixels  $i$ . The **S**-matrix leads to a constant variance on strictly all object pixels; while the **H1** and **W1**-matrices lead to a constant variance on most object pixels, but with differences on few pixels.

To conclude this work, we elaborate on some common characteristics of matrices leading to a strictly constant variance ( $V(\hat{x}_i) = k\bar{x} \forall i$ ). In other words, we seek conditions on the matrix  $\mathbf{A} \in \mathbb{R}_+^{N \times N}$  so that the estimation variance is equal to a constant times the object average  $\bar{x}$ , on every pixel  $i$

$$\forall \mathbf{x} \in \mathbb{R}^+, \quad \mathbf{V}(\hat{\mathbf{x}}) = k\bar{x}\mathbf{1}_N \quad (6.0.1)$$

where  $k \in \mathbb{R}_+^*$  is some constant.

We remind that we consider shot-noise limited measurements  $\mathbf{b} \sim \text{Poisson}(\mathbf{Ax})$ , where  $\mathbf{A} \in \mathbb{R}^{N \times N}$  is an invertible multiplexing matrix with coefficients  $a_{ij} \in \mathbb{R}^+$ .  $\mathbf{x} \in \mathbb{R}^+$  is the intensity object. If the estimate is obtained via least-square estimation ( $\hat{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{b}$ ), the estimation variance reads  $\mathbf{V}(\hat{\mathbf{x}}) = (\mathbf{A}^{-1} \odot \mathbf{A}^{-1})\mathbf{Ax}$ .

### 6.1 A necessary condition

Here,  $\forall \mathbf{x} \in \mathbb{R}^+$ , we seek invertible matrices  $\mathbf{A} \in \mathbb{R}_+^{N^2}$  such that:

$$\begin{aligned} \mathbf{V}(\hat{\mathbf{x}}) &= k\bar{x}\mathbf{1}_N \\ \Leftrightarrow (\mathbf{A}^{-1} \odot \mathbf{A}^{-1})\mathbf{Ax} &= \frac{k}{N}\mathbf{J}_N\mathbf{x} \\ \Leftrightarrow (\mathbf{A}^{-1})^{\odot 2} &= \frac{k}{N}\mathbf{J}_N\mathbf{A}^{-1} \\ \Leftrightarrow t_{ij}^2 &= \frac{k}{N} \sum_q t_{qj} \quad \forall q, j \end{aligned}$$

where  $t_{ij} = (\mathbf{A}^{-1})_{ij}$  denotes the elements of  $\mathbf{A}^{-1}$  and  $\sum_q t_{qj}$  is the sum of the elements of the column  $j$  of  $\mathbf{A}^{-1}$ . This means that all the elements of a given column  $j$  of  $(\mathbf{A}^{-1})^{\odot 2}$  are equal, and proportional to the sum of the elements of the column  $j$  of  $\mathbf{A}^{-1}$ . This also means that the absolute value of all the elements of a given column  $j$  of  $\mathbf{A}^{-1}$  are equal. Defining  $l_j = \sqrt{\frac{k}{N} \sum_q t_{qj}} > 0$  ( $l_j \neq 0$  since  $\mathbf{A}$  is invertible), we have  $t_{ij}^2 = l_j^2$  and  $t_{ij} = \pm l_j = \delta_{ij}l_j$ , with  $\delta_{ij} = -1$  or  $+1$ .

In the matrix form, this reads:

$$(\mathbf{A}^{-1})^{\odot 2} = \begin{pmatrix} l_1^2 & \cdots & l_N^2 \\ \vdots & \vdots & \vdots \\ l_1^2 & \cdots & l_N^2 \end{pmatrix} = \begin{pmatrix} t_{11}^2 & \cdots & t_{1N}^2 \\ \vdots & \ddots & \vdots \\ t_{N1}^2 & \cdots & t_{NN}^2 \end{pmatrix} = \frac{k}{N} \begin{pmatrix} \sum_i t_{i1} & \cdots & \sum_i t_{iN} \\ \vdots & & \vdots \\ \sum_i t_{i1} & \cdots & \sum_i t_{iN} \end{pmatrix} \quad (6.1.2)$$

and therefore:

$$\mathbf{A}^{-1} = \begin{pmatrix} \delta_{11}l_1 & \cdots & \delta_{1N}l_N \\ \vdots & \ddots & \vdots \\ \delta_{1N}l_1 & \cdots & \delta_{NN}l_N \end{pmatrix} \quad (6.1.3)$$

Again, we know that all the elements of a given column  $j$  of  $(\mathbf{A}^{-1})^{\odot 2}$  are equal to  $k/N$  times the sum of the elements of the column  $j$  of  $\mathbf{A}^{-1}$ . Denoting  $p_j$  the number of negative elements (i.e. with  $\delta_{ij} = -1$ ) in column  $j$ , and  $(N - p_j)$  the number of positive elements (i.e. with  $\delta_{ij} = +1$ ) in column  $j$ , we have,  $\forall$  column  $j$ :

$$\begin{aligned} l_j^2 &= \frac{k}{N} ((N - p_j)l_j - p_j l_j) \\ \Leftrightarrow l_j^2 &= \frac{k}{N} (N - 2p_j)l_j \\ \Leftrightarrow l_j &= \frac{k}{N} (N - 2p_j) \quad (l_j \neq 0) \end{aligned}$$

Note that since  $l_j > 0$ , there must be strictly more positive than negative elements in each column of  $\mathbf{A}^{-1}$ : if  $N$  is odd, the maximum number of negative elements in column  $j$  is  $\frac{N-1}{2}$ ; if  $N$  is even, the maximum number of negative elements in column  $j$  is  $\frac{N}{2} - 1$ . Last, since  $\mathbf{A}$  is invertible,  $p_j = 0$  in one column at most, and therefore  $N > 2$ .

Overall, we showed that,  $\forall \mathbf{x} \in \mathbb{R}^+$ ,  $\forall N > 2$ , and for  $\mathbf{A} \in \mathbb{R}_+^{N^2}$  invertible:

$$\mathbf{V}(\hat{\mathbf{x}}) = k\bar{x}\mathbf{1}_N \implies \mathbf{A}^{-1} = \frac{k}{N} \begin{pmatrix} \delta_{11}(N - 2p_1) & \cdots & \delta_{1N}(N - 2p_N) \\ \vdots & \vdots & \vdots \\ \delta_{1N}(N - 2p_1) & \cdots & \delta_{NN}(N - 2p_N) \end{pmatrix} \quad (6.1.4)$$

where  $k \in \mathbb{R}^{+*}$  is some constant,  $\delta_{ij} = -1$  or  $+1$ , and  $p_j$  is the number of negative elements  $\delta_{ij} = -1$  in column  $j$  (with  $p_j \leq \frac{N-1}{2}$  if  $N$  is odd,  $p_j \leq \frac{N}{2} - 1$  if  $N$  even, and  $p_j = 0$  in one column at most). The sum of the column  $j$  of  $\mathbf{A}^{-1}$  is  $(k/N)(N - 2p_j)^2$ . In addition, no column can be a linear combination of the other, which limits the number of possibilities.

In conclusion,  $\mathbf{A}^{-1}$  is therefore a matrix in which the absolute value of all elements of a given column  $j$  are equal: it is equal to a number that only depends on the number of negative element of the same column. Note that this condition is only necessary and not sufficient.

Note on special matrices:

The  $\mathbf{S}$ -matrix used throughout this study verifies (6.1.4) with,  $\forall j$ ,  $p_j = p = \frac{N-1}{2}$ ,  $l_j = l = \frac{k}{N}$  and  $k = \frac{2N}{N+1} \approx 2$ .

Note that  $\mathbf{H1}$  does not does not verify (6.1.4) and does not exactly lead to  $\mathbf{V}(\hat{\mathbf{x}}) = k\bar{x}\mathbf{1}_N$ , in particular due to its first row and column.

## 6.2 Sufficient condition 1

Here, we consider  $\mathbf{A}^{-1}$  of the form of (6.1.4), and add the hypothesis that all  $p_j$  are equal, i.e. that,  $\forall j$ :

$$p_j = p, \text{ i.e. } l_j = \frac{k}{N}(N - 2p) \quad (6.2.5)$$

Then, we have:

$$\mathbf{A}^{-1} = \frac{k}{N}(N - 2p) \begin{pmatrix} \delta_{11} & \cdots & \delta_{1N} \\ \vdots & \ddots & \vdots \\ \delta_{1N} & \cdots & \delta_{NN} \end{pmatrix} = \frac{k}{N}(N - 2p)\mathbf{\Lambda} \quad (6.2.6)$$

Therefore,  $\mathbf{A}^{-1}$  is proportional to a matrix  $\mathbf{\Lambda}$  made of  $+1$  and  $-1$  signs ( $|\delta_{ij}| = 1$ ), where there are exactly  $p$  negative elements in each column. Note that many combinations of signs, in particular permutations, may be found. Note also that the sum of all columns of  $\mathbf{A}^{-1}$  is a constant:

$$\mathbf{J}_N \mathbf{A}^{-1} = \frac{k}{N}(N - 2p)^2 \mathbf{J}_N \quad (6.2.7)$$

and therefore the sum of all columns of  $\mathbf{A}$  is also a constant:

$$\mathbf{J}_N \mathbf{A} = \frac{N}{k(N - 2p)^2} \mathbf{J}_N \quad (6.2.8)$$

Then, it is easy to verify that:

$$\mathbf{V}(\hat{\mathbf{x}}) = (\mathbf{A}^{-1})^{\odot 2} \mathbf{A} \mathbf{x} = \frac{k^2}{N^2}(N - 2p)^2 \mathbf{J}_N \mathbf{A} \mathbf{x} = \frac{k^2}{N^2}(N - 2p)^2 \frac{N}{k(N - 2p)^2} \mathbf{J}_N \mathbf{x} = \frac{k}{N} \mathbf{J}_N \mathbf{x} = k\bar{x} \mathbf{1}_N$$

Therefore, we can write,  $\forall \mathbf{x} \in \mathbb{R}^+$ ,  $\forall N > 2$ , and for  $\mathbf{A} \in \mathbb{R}_+^{N^2}$  invertible:

$$\boxed{\mathbf{A}^{-1} = \frac{k}{N}(N - 2p)\mathbf{\Lambda} \implies \mathbf{V}(\hat{\mathbf{x}}) = k\bar{x} \mathbf{1}_N} \quad (6.2.9)$$

where  $k \in \mathbb{R}_+^*$  is some constant and  $\mathbf{\Lambda}$  is a matrix with elements  $\delta_{ij} = \pm 1$ , in which each column contains  $p$  negative elements.

Note on special matrices:

The  $\mathbf{S}$ -matrix used throughout this study verifies (6.2.9) with  $p = \frac{N-1}{2}$ ,  $l = \frac{k}{N}$  and  $k = \frac{2N}{N+1} \approx 2$ . Note that  $\mathbf{H1}$  does not verify (6.2.9) in particular due to its first row and column.

## 6.3 Sufficient condition 2

Independently from the above results, another sufficient condition is straightforward. Since  $\mathbf{V}(\hat{\mathbf{x}}) = (\mathbf{A}^{-1} \odot \mathbf{A}^{-1}) \mathbf{A} \mathbf{x}$ , an invertible matrix  $\mathbf{A} \in \mathbb{R}_+^{N^2}$  in which the sum of its columns is constant and the Hadamard product of its inverse is proportional to  $\mathbf{J}_N$  leads to a variance equal to a constant times the object average  $\bar{x}$ :

$$\begin{cases} \mathbf{A}^{-1} \odot \mathbf{A}^{-1} = \frac{\theta_1}{N} \mathbf{J}_N \\ \mathbf{J}_N \mathbf{A} = \theta_2 \mathbf{J}_N \end{cases} \implies \mathbf{V}(\hat{\mathbf{x}}) = \theta_1 \theta_2 \bar{x} \mathbf{1}_N = k\bar{x} \mathbf{1}_N \quad (6.3.10)$$

where  $\theta_1 \in \mathbb{R}^{+*}$  and  $\theta_2 \in \mathbb{R}^+$  are some constants.

# Bibliography

- [1] Matthew P. Edgar, Graham M. Gibson, and Miles J. Padgett. Principles and prospects for single-pixel imaging. *Nature Photonics*, 13(1):13–20, 2019.
- [2] P. Fellgett. Conclusions on Multiplex Methods. *Le Journal de Physique Colloques*, 28(C2):C2–165–C2–171, 1967.
- [3] R. A. DeVerse, R. M. Hammaker, and W. G. Fateley. Realization of the Hadamard multiplex advantage using a programmable optical mask in a dispersive flat-field near-infrared spectrometer. *Applied Spectroscopy*, 54(12):1751–1758, 2000.
- [4] Kaare Brandt Petersen and Michael Syskind Pedersen. *The Matrix Cookbook*. 2012.
- [5] A. D. Rodríguez, P. Clemente, E. Tajahuerce, and J. Lancis. Dual-mode optical microscope based on single-pixel imaging. *Optics and Lasers in Engineering*, 82:87–94, 2016.
- [6] F. Soldevila, P. Clemente, E. Tajahuerce, N. Uribe-Patarroyo, P. Andrés, and J. Lancis. Computational imaging with a balanced detector. *Scientific Reports*, 6(March):1–10, 2016.
- [7] Zibang Zhang, Xiao Ma, and Jingang Zhong. Single-pixel imaging by means of Fourier spectrum acquisition. *Nature Communications*, 6(September 2017):1–6, 2015.
- [8] Yiwei Zhang, Graham M. Gibson, Matthew P. Edgar, Giles Hammond, and Miles J. Padgett. Dual-band single-pixel telescope. *Optics Express*, 28(12):18180–18188, 2020.
- [9] Martin Harwit, Neil J A Sloane, I Sloane, and Neil James. *Hadamard Transform Optics*. Academic Press, 1979.
- [10] Vincent Studer, Jérôme Bobin, Makhlad Chahid, Hamed Shams, and Emmanuel Candes. Compressive fluorescence microscopy for biological and hyperspectral imaging. *PNAS*, 109(26):1679–1687, 2012.
- [11] Pascal Berto, Camille Scotté, Frédéric Galland, Hervé Rigneault, and Hilton B. de Aguiar. Programmable single-pixel-based broadband stimulated Raman scattering. *Optics Letters*, 42(9):1696, 2017.
- [12] Camille Scotté, Siddharth Sivankutty, Randy A. Bartels, and Hervé Rigneault. Line-scan compressive Raman imaging with spatiospectral encoding. *Optics Letters*, 45(19):5567, 2020.
- [13] Camille Scotté. *Spontaneous Compressive Raman technology : developments and applications*. PhD thesis, Aix-Marseille University, 2020.
- [14] J.J. Sylvester. Thoughts on inverse orthogonal matrices, simultaneous sign-successions, and tessellated pavements in two or more colours, with applications to Newton’s rule, ornamental tile-work, and the theory of numbers. *Philosophical Magazine*, 34:461–495, 1867.
- [15] A. Hedayat and W.D Wallis. Hadamard matrices and their applications. *The Annals of Statistics*, 6(6):1184–1238, 1978.
- [16] Jennifer Seberry Wallis. On the existence of Hadamard matrices. *Journal of Combinatorial Theory, Series A*, 21(2):188–195, 1976.

- [17] Ching-Shui Cheng. An Application of the Kiefer-Wolfowitz Equivalence Theorem to a Problem in Hadamard Transform Optics. *The Annals of Statistics*, 15(4):1593–1603, 1987.
- [18] Roman Drnovšek. On the S-matrix conjecture. pages 1–6, 2013.
- [19] Greg Futia, Philip Schlup, David G Winters, and Randy a Bartels. Spatially-chirped modulation imaging of absorption and fluorescent objects on single-element optical detector. *Optics express*, 19(2):1626–1640, 2011.
- [20] Zibang Zhang, Xueying Wang, Guoan Zheng, and Jingang Zhong. Hadamard single-pixel imaging versus Fourier single-pixel imaging. *Optics Express*, 25(16):19619–19639, 2017.
- [21] Zibang Zhang, Xueying Wang, Guoan Zheng, and Jingang Zhong. Fast Fourier single-pixel imaging via binary illumination. *Scientific Reports*, 7(1):1–9, 2017.
- [22] Ling Tong Meng, Ping Jia, Hong Hai Shen, Ming Jie Sun, Dong Yao, Han Yu Wang, and Chun Hui Yan. Sinusoidal single-pixel imaging based on fourier positive–negative intensity correlation. *Sensors*, 20(6), 2020.
- [23] A Moshtaghpour and L Jacques. Compressive Hyperspectral Imaging : Fourier Transform Interferometry meets Single Pixel Camera. pages 1–5, 2018.
- [24] Gilbert Strang. The discrete cosine transform. *SIAM Review*, 41(1):135–147, 1999.
- [25] Peter Fellgett. *Theory of infrared sensitivities and its application to investigations of stellar radiation in the near infrared*. PhD thesis, University of Cambridge, 1951.
- [26] Daniel R Fuhrmann, Chrysanthe Preza, Joseph A O’sullivan, Donald L Snyder, and William H Smith. Spectrum Estimation from Quantum-Limited Interferograms. *IEEE Transactions on Signal Processing*, 52(4):950–961, 2004.
- [27] Camille Scotté, Siddharth Sivankutty, Patrick Stockton, Randy A. Bartels, and Hervé Rigneault. Compressive Raman imaging with spatial frequency modulated illumination. *Optics Letters*, 44(8):1936–1939, 2019.
- [28] Richard R Treffers. Signal-to-noise ratio in Fourier spectroscopy. *Applied Optics*, 16(12):3103, 1977.
- [29] Wojciech Tadej and Karol Życzkowski. A concise guide to complex Hadamard matrices. *Open Systems and Information Dynamics*, 13(2):133–177, 2006.
- [30] Teo Banica. Complex Hadamard matrices and applications. 2019.