Optimal multiplexed sensing: bounds, conditions and a graph theory link

Netanel Ratner,[1] Yoav Y. Schechner,[1,*] and Felix Goldberg[2]

[1]*Dept. Electrical Engineering, Technion - Israel Inst. Technology*
Haifa 32000, Israel

[2]*Dept. Mathematics, Technion - Israel Inst. Technology*
Haifa 32000, Israel

[*]*Corresponding author: yoav@ee.technion.ac.il*

Abstract: Measuring an array of variables is central to many systems, including imagers (array of pixels), spectrometers (array of spectral bands) and lighting systems. Each of the measurements, however, is prone to noise and potential sensor saturation. It is recognized by a growing number of methods that such problems can be reduced by multiplexing the measured variables. In each measurement, multiple variables (radiation channels) are mixed (multiplexed) by a code. Then, after data acquisition, the variables are decoupled computationally in post processing. Potential benefits of the use of multiplexing include increased signal-to-noise ratio and accommodation of scene dynamic range. However, existing multiplexing schemes, including Hadamard-based codes, are inhibited by fundamental limits set by sensor saturation and Poisson distributed photon noise, which is scene dependent. There is thus a need to find optimal codes that best increase the signal to noise ratio, while accounting for these effects. Hence, this paper deals with the pursuit of such optimal measurements that avoid saturation and account for the signal dependency of noise. The paper derives lower bounds on the mean square error of demultiplexed variables. This is useful for assessing the optimality of numerically-searched multiplexing codes, thus expediting the numerical search. Furthermore, the paper states the necessary conditions for attaining the lower bounds by a general code. We show that *graph theory* can be harnessed for finding such ideal codes, by the use of strongly regular graphs.

© 2007 Optical Society of America

References and links

1. D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk. "A new compressive imaging camera architecture using optical-domain compression." In Proc. SPIE volume **6065** (2006).
2. W. G. Fateley, R. M. Hammaker, R. A. DeVerse, R. R. Coifman, and F. B. Geshwind. "The other spectroscopy: demonstration of a new de-dispersion imaging spectrograph." Vib. Spectrosc. **29**:163–170 (2002).
3. C. Fernandez, B. D. Guenther, M. E. Gehm, D. J. Brady, and M. E. Sullivan. "Longwave infrared (LWIR) coded aperture dispersive spectrometer." Opt. Express **15**:5742–5753 (2007).

4. M. E. Gehm, S. T. McCain, N. P. Pitsianis, D. J. Brady, P. Potuluri, and M. E. Sullivan. "Static two-dimensional aperture coding for multimodal, multiplex spectroscopy." Appl. Opt. **43**:2965–2974 (2006).

5. Q. S. Hanley, D. J. Arndt-Jovin, and T. M. Jovin. "Spectrally resolved fluorescence lifetime imaging microscopy." Appl. Spectrosc. **56**:63–84 (2002).

6. G. Nitzsche and R. Riesenberg. "Noise, fluctuation and HADAMARD-transform-spectrometry." In Proc. SPIE volume **5111**, pages 273–282 (2003).

7. J. F. Turner and P. J. Treado. "Adaptive filtering and hadamard transform imaging spectroscopy with an acousto-optic tunable filter (AOTF)." In Proc. SPIE volume **2599**, pages 285–293 (1996).

8. E. E. Fenimore and T. M. Cannon. "Coded aparture imaging with uniformly redundant arrays." Appl. Opt. **17**:337–347 (1978).

9. M. Harwit and N. J. A. Sloane. *Hadamard Transform Optics.* Academic Press, New York (1979).

10. T. M. Palmieri. "Multiplex methods and advantages in X-ray astronomy." Astrophysics and Space Science **28**:277–287 (1974).

11. R. J. Proctor, G. K. Skinner, and A. P. Willmore. "The design of optimum coded mask X-ray telescopes." Royal Astronomical Society, Monthly Notices **187**:633–643 (1979).

12. G. K. Skinner. "X-ray imaging with coded masks." Scientific American **259**:84–89 (1988).

13. A. M. Bronstein, M. M. Bronstein, E. Gordon, and R. Kimmel. "Fusion of 2d and 3d data in three-dimensional face recognition." In Proc. IEEE ICIP Vol. **1**, pages 87–90 (2004).

14. O. G. Cula, K. J. Dana, D. K. Pai, and D. Wang. "Polarization multiplexing and demultiplexing for appearance-based modeling." IEEE Trans. PAMI **29**:362–367 (2007).

15. K. C. Lee, J. Ho, and D. J. Kriegman. "Acquiring linear subspaces for face recognition under variable lighting." IEEE Trans. PAMI **27**:684–698 (2005).

16. M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas. "Synthetic aperture confocal imaging." ACM TOG **23**:825–834 (2004).

17. F. Moreno-Noguer, S. K.Nayar, and P. N. Belhumeur. "Optimal illumination for image and video relighting." In Proc. CVMP pages 201–210 (2005).

18. Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. "A theory of multiplexed illumination." In Proc. IEEE ICCV Vol. **2**, pages 808–815 (2003).

19. A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec. "Performance relighting and reflectance transformation with time-multiplexed illumination." ACM TOG **24**:756–764 (2005).

20. A. Busboom, H. D. Schotten, and H. Elders-Boll. "Coded aperture imaging with multiple measurements." J. Opt. Soc. Am. A **14**(5):1058–1065 (1997).

21. E. E. Fenimore. "Coded aperture imaging: predicted performance of uniformly redundant arrays." Appl. Opt. **17**:3562–3570 (1978).

22. N. Ratner and Y. Y. Schechner. "Illumination multiplexing within fundamental limits." In Proc. IEEE CVPR (2007).

23. A. Wuttig. "Optimal transformations for optical multiplex measurements in the presence of photon noise." Appl. Opt. **44**:2710–2719 (2005).

24. Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. "Multiplexing for optimal lighting." IEEE Trans. PAMI **29**:1339–1354 (2007).

25. V. P. Kozlov and E. V. Sedunov. "Optimization of multiplex measuring systems in the presence of statistical signal fluctuations." Cybernetics and Systems Analysis **28**:830–839 (1992).

26. Y. A. Shutova. "Optimization of binary masks for Hadamard-transform optical spectrometers". J. Opt. Technol. **67**:50–53 (2000).

27. C. D. Meyer. *Matrix Analysis and Applied Linear Algebra.* SIAM (2000).

28. R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge, New York (1985).

29. M. T. Chu. "A fast recursive algorithm for constructing matrices with prescribed eigenvalues and singular values." SIAM J. on Numerical Analysis **37**(3):1004–1020 (2000).

30. R. Diestel. *Graph Theory.* Springer, 3rd edition (2000).

31. P. J. Cameron and J. H. V. Lint. *Designs, Graphs, Codes, and Their Links.* Cambridge University Press, New York, NY, USA (1991).

32. J. J. Seidel. "Strongly regular graphs with (-1, 1, 0) adjacency matrix having eigenvalue 3." Linear Algebra Appl. **1**:281–289 (1968).

33. W. Haemers. "Matrix techniques for strongly regular graphs and related geometries." Intensive Course on Finite Geometry and its Applications, University of Ghent (2000).

34. M. Alicacute, B. Mond, J. Pecbreve aricacute and V. Volenec. "The arithmetic-geometric-harmonic-mean and related matrix inequalities." Linear Algebra and its Applications **264**(1):55–62 (1997).

35. K. Coolsaet and J. Degraer. "The strongly regular (45,12,3,3) graphs." Elec. Journ. Combin **13**(1) (2006).

36. T. Spence. "Strongly Regular Graphs on at most 64 vertices" http://www.maths.gla.ac.uk/ es/srgraphs.html

37. G. Royle. "Strongly regular graphs" (1996) http://people.csse.uwa.edu.au/gordon/remote/srgs/index.html

38. P. Puxley and T. Geballe. "Transmission Spectra" (1999) http://www.gemini.edu/sciops/ObsProcess/obsConstraints/ocTransSpectra.html

39. S. Ioué and K. R. Spring. *Video Microscopy*, 2nd ed. ch. 6,7,8, Plenum Press, New York. (1997).
40. C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang. "Noise estimation from a single image." In Proc. CVPR Vol. **1** pages 901–908 (2006).
41. F. Alter, Y. Matsushita, and X. Tang. "An intensity similarity measure in low-light conditions." In Proc. ECCV Vol. **4**, pages 267–280 (2006).
42. H. H. Barrett and W. Swindell. *Radiological Imaging*, volume **1**. Academic press, New York (1981).

1. Introduction

Often, there is a need to measure an array of variables. This occurs in two dimensional imaging [1] (array of pixels), spectrometry (array of wavelength bands) [2–7], tomography and coded aperture imaging (array of viewing directions) [8–12], and study of reflectance functions in computer vision (array of lighting directions) [13–19] etc. The array can be high dimensional, if it includes several types of the mentioned domains, such as in hyperspectral imaging. The variables in the array are often measured sequentially. For example, in each measurement, the lighting direction, spectral band, or an aperture mask can change.

Each of the measurements is prone to noise. This measurement noise propagates to an error in the estimation of the desired array of variables. This problem is exacerbated by constraints on the system and the object. For example, according to Ref. [3], constrains on system weight, price and power consumption of infrared spectrometers may require the use of uncooled detectors, which are generally noisy. In another example, the radiating object may be wide and diffuse [4], making its coupling into a narrow slit of a spectrometer inefficient, thus yielding a low signal, relative to the noise. For a given acquisition time, however, noise can be efficiently reduced using *multiplexing* [2–7,9,18–23]. Here, in each sequential measurement, multiple elements of the sought array (e.g., multiple spectral bands) are linearly weighted (multiplexed) and sensed simultaneously at a detector. In sequential measurements, the multiplex weighting changes. The set of weights is termed a multiplexing *code*. Once the measurements are done, computational demultiplexing derives the sought variable array.

This paper deals with pursuit of an optimal code, such that the noise in the sought variables is minimal. Common multiplexing codes are based on Hadamard matrices [9]. Their use was proposed in a wide range of fields, including very recent studies [3, 4, 18, 19, 24]. They are optimal if noise is independent of the signal and if no saturation effect is possible. Hence, they do not account for two fundamental effects: *saturation* and *photon noise* (the latter is signal-dependent). This can make multiplexing by Hadamard codes counterproductive [19, 24], and may be the reason for low benefits achieved in practice by Hadamard-based systems [3, 4, 18, 19, 24]. Moreover, for most sizes of the variable array, there are no Hadamard codes. In contrast, in this paper we deal with general codes, considering explicitly saturation and signal-dependent photon noise. We seek to maximize the estimation accuracy of the signal sources, by accounting for the sensor specifications and avoiding saturation.

Recent studies [20, 22, 23, 25, 26] considered some of these fundamental aspects. Ref. [20] derived an expression for the signal to noise ratio (SNR) in the presence of photon noise. It utilized this expression to examine the SNR yielded by pre-set cyclic codes. Ref. [23] considered photon-noise for devising binary, cyclic codes, for a small set of sizes of the sought variable array. To account for any array size, saturation and photo noise, Ref. [22] proposed efficient numerical optimization of the multiplexing code. However, numerical optimization as in Ref. [22] may stagnate at sub-optimal solutions, with no indication of how good they are relative to the elusive optimal solution.

We seek to properly understand the task of optimal multiplexing in general, and even point out solutions to the task. Thus, this paper performs a comprehensive theoretical analysis. First, it focuses on the search for optimal codes, under fixed radiant power, as on a saturation threshold. It derives fundamental *lower bounds* on the output noise (estimation error). These bounds

indicates how well any general multiplexing code can potentially reduce the noise. Furthermore, we show that in order to attain a lower bound (*ideal* multiplexing), certain conditions should be satisfied by the multiplexing code. Hadamard codes are special cases of our analysis, hence this work significantly generalizes the known art. This has a major importance, as we explain a way to numerically optimize multiplexing codes under the influence of photo noise, beside saturation and detector noise.

In addition, an interesting relation is revealed to *graph theory*, a large mathematical field, typically distinct from that used in intensity noise analysis. We show that graph theory can be harnessed for finding ideal codes, by using *strongly regular graphs*.

2. Theoretical background

2.1. Multiplexing

As mentioned in Sec. 1, multiplexing is a general method used in different fields. In each field, the sought (demultiplexed) array of variables represents a different physical quantity. The same applies to the multiplexed measurements. Nevertheless, the mathematical treatment is equivalent in all the fields. To associate the variables and measurements with familiar quantities, we use the domain of lighting, which is employed in computer vision [13–15,18] and in research done in computer graphics [16, 17, 19]. However, thanks to the equivalence of the algebra, analogies can be easily made to the other fields, such as spectrometry and hyperspectral imaging.

In the field of lighting, an object is viewed and measured from a fixed location. In each measurement, the object is irradiated from a different direction, by a radiation source. Under this irradiance, the object is imaged. Consider a setup of N radiation sources. Let $\mathbf{i} = (i_1, i_2, \ldots, i_N)^t$ be the set of image irradiance values, at a certain pixel. Each element in $\mathbf{i}$ corresponds to irradiance by any *individual* source in this setup. Here t denotes transposition.

In general, several sources can be turned on in each measurement (multiplexing). Define an $N \times N$ multiplexing matrix $\mathbf{W}$, referred to as a *multiplexing code*. Each element of its m'th row represents the normalized radiance of a source in the m'th measurement. The radiance of a source is normalized relative to its maximum value. Hence, an element value of 0 states that the source is completely off, while 1 indicates a fully activated source. The sequential measurements acquired at a detector are denoted by the vector $\mathbf{a} = (a_1, a_2, \ldots, a_N)^t$, given by

$$\mathbf{a} = \mathbf{W}\mathbf{i} + \upsilon \ . \tag{1}$$

Here υ is the measurement noise. Any bias to this noise is assumed to be compensated for. Let the noise υ be uncorrelated in different measurements, and let its variance be σ_a^2.

Once measurements have been acquired under multiplexed sources, those measurements are de-multiplexed computationally. This derives estimates for the irradiance values of the object under the individual radiation sources $\hat{\mathbf{i}}$. The best unbiased linear estimator in the sense of mean square error (MSE) for the irradiance corresponding to the individual-sources is

$$\hat{\mathbf{i}} = \mathbf{W}^{-1}\mathbf{a} \ . \tag{2}$$

The MSE of this estimator [9, 18] is

$$\mathrm{MSE}_{\hat{\mathbf{i}}} = \frac{\sigma_a^2}{N} \mathrm{trace}\left[\left(\mathbf{W}^t\mathbf{W}\right)^{-1}\right] \ . \tag{3}$$

This is the expected noise variance of the recovered sources. In this paper we pursue the problem of finding a multiplexing code $\mathbf{W}$ that maximizes the SNR of $\hat{\mathbf{i}}$, or, equivalently, minimizes $\mathrm{MSE}_{\hat{\mathbf{i}}}$. Specifically, we seek a lower bound on Eq. (3) and derive conditions on $\mathbf{W}$ to attain this bound, hence minimizing $\mathrm{MSE}_{\hat{\mathbf{i}}}$.

2.2. Eigenvalues and singular values

In this section we briefly review elementary definitions and results from linear algebra that will later be used for our analysis.

Definition Let $\Lambda = \{\lambda_f\}_{f=1}^{N}$ be the set of the eigenvalues (EVs) of a matrix $\mathbf{W}$. The *multiplicity* of $\lambda_f \in \Lambda$ is the number of repetitions of the value λ_f in Λ.

Lemma 1. *If $\mathbf{R}$ and $\mathbf{Q}$ are matrices such that $\mathbf{RQ}$ is a square matrix, then* [27]

$$\mathrm{trace}(\mathbf{RQ}) = \mathrm{trace}(\mathbf{QR}) \quad . \tag{4}$$

Lemma 2. *Let $\mathbf{W}$ be a non-singular $N \times N$ matrix. Its EVs are $\lambda_1 \leqslant \ldots \leqslant \lambda_N$. Then (See for example Ref.* [28]*)*
i)

$$\mathrm{trace}(\mathbf{W}) = \sum_{f=1}^{N} \lambda_f \quad . \tag{5}$$

ii) The EVs of $\mathbf{W}^{-1}$ are $\lambda_N^{-1} \leqslant \ldots \leqslant \lambda_1^{-1}$.

Definition Let $\mu_1 \leqslant \mu_2 \leqslant \ldots \leqslant \mu_N$ be the EVs of $\mathbf{W}^t\mathbf{W}$. Then, the singular values (SVs) of $\mathbf{W}$, $\{\xi_f\}_{f=1}^{N}$ are defined as

$$\xi_f = \sqrt{\mu_f} \quad . \tag{6}$$

Note that if $\mathbf{W}$ is symmetric, then $\mathbf{W}^t\mathbf{W} = \mathbf{W}^2$ and

$$\xi_m = |\lambda_m| \quad \forall m \quad . \tag{7}$$

Ref. [29] quotes the following theorem.

Theorem 1. *(Weyl-Horn)*

$$\prod_{m=f}^{N} \xi_m \geqslant \prod_{m=f}^{N} |\lambda_m| \quad \forall f \in \{2,\ldots,N\} \tag{8}$$

and

$$\prod_{m=1}^{N} \xi_m = \prod_{m=1}^{N} |\lambda_m| \quad . \tag{9}$$

2.3. Strongly regular graphs

We now refer to some elementary definitions from graph theory. We will use them when seeking optimal solutions to the multiplexing problem. We quote some basic definitions from Ref. [30].

Consider a graph $G = (V,E)$, where V is a set of N *vertices*. Here E is the set of *edges* connecting a pair of vertices.

Definition Two vertices p,q are said to be *adjacent* or *neighbors* if they are connected by an edge.

Definition The $N \times N$ *adjacency* matrix Ω of the graph G is composed of elements

$$\omega_{p,q} = \begin{cases} 1 & \text{if } p \text{ and } q \text{ are neighbors} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$
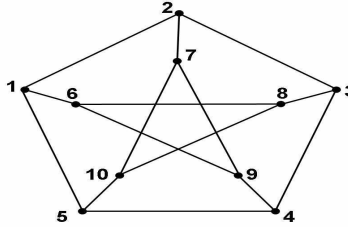
Fig. 1. An example of a strongly regular graph (Peterson) [31]. This graph has the parameters $(N = 10; k = 3; \alpha = 0; \beta = 1)$.

Definition The *complement* of a graph G is a graph $\bar{G}$ where its adjacency matrix of $\bar{\Omega}$, is composed of elements

$$\bar{\omega}_{p,q} = \begin{cases} 1 & \text{if } \omega_{p,q} = 0 \text{ and } p \neq q \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Definition If all the vertices of G have the same number of neighbors k, then G is *k-regular*. In this case

$$\sum_{q=1}^{N} \omega_{p,q} = k \qquad \forall p \quad . \tag{12}$$

Definition A *strongly regular graph* (SRG) [31] with parameters $(N; k; \alpha; \beta)$ is a *k*-regular graph that has the following properties:

- Any two adjacent vertices have exactly α common neighbors (neighbors of both vertices).

- Any two non-adjacent vertices have exactly β common neighbors.

For example, consider the graph in Fig. 1. The adjacent vertices 5 and 10 have no common neighbors and this relation also applies to all the other adjacent pairs in the graph. Hence, here $\alpha = 0$. Moreover, vertices 5 and 3 have a single common neighbor, 9, and so are all other analogous pairs. Hence, here $\beta = 1$.

Theorem 2. *The parameters $(N; k; \alpha; \beta)$ of a strongly regular graph satisfy* [31] *the constraint*

$$k(k - \alpha - 1) = (N - k - 1)\beta . \tag{13}$$

In the following, we make use of a theorem due to Seidel [32]:

Theorem 3. *Let G be an SRG with parameters $(N; k; \alpha; \beta)$. Its adjacency matrix Ω has generally three distinct EVs,*

$$\lambda_1^{\Omega} = \frac{(\alpha - \beta) + \sqrt{\Delta}}{2} \tag{14}$$

$$\lambda_2^{\Omega} = \frac{(\alpha - \beta) - \sqrt{\Delta}}{2} \tag{15}$$

$$\lambda_3^{\Omega} = k \quad , \tag{16}$$

where,

$$\Delta \equiv (\alpha - \beta)^2 + 4(k - \beta) \quad . \tag{17}$$

The multiplicity of λ_3^Ω is 1.

From Eq. (10), Ω is symmetric. Thus, Eq. (7) applies to the SVs of Ω.

$$\xi_f^\Omega = |\lambda_f^\Omega| \quad \forall f \quad . \tag{18}$$

Since the EVs of Ω indicate its SVs, Theorem 3 can be applied to the SVs of Ω. In particular, the multiplicity of EVs in Theorem 3 generally applies to the SVs of Ω.

3. Optimal power-regulated multiplexing

3.1. Problem formulation

We now seek multiplexing codes that minimize $\mathrm{MSE}_{\hat{\mathbf{i}}}$ under the constraint of a fixed irradiance of an object in each measurement. Such a constraint is desired to avoid saturation of the detector. In saturation, the number of electrons generated in the detector exceeds the capacity of the circuitry that translates it into gray levels. The property of a fixed power is useful for other reasons, such as curbing photon noise, as we shall detail later.

To better express the fixed irradiance constraint, we define the variable C. It is the *effective number of radiation sources* used in each measurement. This unit-less variable is equivalent to the total radiant power used in a measurement. For example, in Hadamard-based codes,

$$C = C_{\mathrm{Had}} \equiv \frac{N+1}{2} \quad . \tag{19}$$

Fixing the scalar C is equivalent to restricting the power emitted jointly by all sources. Suppose that a multiplexing code $\mathbf{W}$ is suggested, such that the irradiance by all the sources exceeds the pre-set threshold. Then this code cannot be used: it yields to much radiation. For example, in a system where saturation occurs if C is greater than $C_{\mathrm{sat}} = N/5$, Hadamard multiplexing cannot be used, since C_{Had} exceeds C_{sat} in this case.

If $\mathbf{W}$ violates the fixed-power constraint, what can be done? A trivial way is to equally reduce the power of each of the sources. However, refs. [18, 22] proved that such a step should be avoided. A better solution is to modify $\mathbf{W}$, such that its corresponding C would comply with the constraint. Such a modification, however, is not easy to accomplish. The reason is that current codes in the literature [9, 23] do not support multiplexing of an arbitrary number C out of N sources. Specifically, Hadamard codes are too limited: according to Eq. (19), these codes totally couple C to N. This raises the need to extend the set of multiplexing codes, to comply with a general constraint on the effective number of simultaneously activated sources C.

Power is fixed by setting

$$\sum_{s=1}^{N} w_{m,s} = C \quad \forall m \in \{1, 2, \ldots, N\} \quad . \tag{20}$$

Recall that each source can be activated with some portion of its maximum power, *i.e.*

$$0 \leqslant w_{m,s} \leqslant 1 \quad \forall m, s \in \{1, 2, \ldots, N\} \quad . \tag{21}$$

We use Eq. (3) to formulate a minimization task of $\mathrm{MSE}_{\hat{\mathbf{i}}}$. To simplify the analysis, we slightly modify the problem for the moment, and define the cost function

$$\widetilde{\mathrm{MSE}} \triangleq \frac{\mathrm{MSE}_{\hat{\mathbf{i}}}}{\sigma_a^2} = \frac{1}{N} \mathrm{trace}\left[\left(\mathbf{W}^t \mathbf{W}\right)^{-1}\right] \quad . \tag{22}$$

Minimizing $\widetilde{\text{MSE}}$ is equivalent to minimizing $\text{MSE}_{\hat{\mathbf{i}}}$, if σ_a^2 is constant. The influence of σ_a^2 will be discussed in Sec. 7.

The constraints for our problem are taken from Eqs. (20,21). Thus, the optimization problem is

$$\min_{\mathbf{W}} \widetilde{\text{MSE}} \triangleq \min_{\mathbf{W}} \frac{1}{N}\text{trace}\left[\left(\mathbf{W}^t\mathbf{W}\right)^{-1}\right] \tag{23}$$

$$\text{s.t.} \quad \sum_{s=1}^{N} w_{m,s} - C = 0 \quad \forall m \in \{1,\ldots,N\} \tag{24}$$

$$-w_{m,s} \leqslant 0 \quad \forall m,s \in \{1,\ldots,N\} \tag{25}$$

$$w_{m,s} - 1 \leqslant 0 \quad \forall m,s \in \{1,\ldots,N\} \ . \tag{26}$$

We shall now derive sufficient conditions for a matrix $\mathbf{W}$ to solve Eqs. (23,24,25,26).

3.2. Conditions for a global optimum

A numerical procedure has been tailored to the optimization problem (23) in Ref. [22]. It is preferable however, to reach a closed-form solution, if it exists. This is done by deriving sufficient conditions for the optimality of a given $\mathbf{W}$. Such conditions allow us to identify an optimal solution to the problem, if a potential solution is encountered. Indeed, later on in Sec. 6 we show that these conditions are satisfied by matrices $\mathbf{W}$ originally developed in graph theory. We can also apply these conditions to verify if a matrix reached by numerical optimization is indeed the global optimum.

Our approach for deriving the optimality conditions is as follows: first, we find a tight lower bound on $\widetilde{\text{MSE}}$. Then, we formulate a necessary and sufficient condition to attain this bound. Finally, we minimize the bound itself, with respect to the elements of $\mathbf{W}$.

3.2.1. The cost as a function of singular values

First, we express Eq. (23) in terms of the SVs of $\mathbf{W}$. Recall from definition 2.2 that $\mu_1 \leqslant \ldots \leqslant \mu_N$ are the EVs of $\mathbf{W}^t\mathbf{W}$. Then using, Lemma 2,

$$\widetilde{\text{MSE}} \triangleq \frac{1}{N}\text{trace}\left[\left(\mathbf{W}^t\mathbf{W}\right)^{-1}\right] = \frac{1}{N}\sum_{f=1}^{N}\frac{1}{\mu_f} \ , \tag{27}$$

Thus, in light of Eq. (6).

$$\mu_f \equiv \xi_f^2 \quad \forall f \in \{1,\ldots,N\} \quad . \tag{28}$$

We show an implication of constraints (24,25,26) on the SVs of $\mathbf{W}$. This will allow us to better streamline (24,25,26) into Eq. (27), forming a lower bound on $\widetilde{\text{MSE}}$. To understand the connection between these constraints and the SVs of $\mathbf{W}$, we cite the following theorem [33]:

Theorem 4. *Let $\mathbf{W}$ be constrained by Eqs. (24,25,26). Then, C is an EV of $\mathbf{W}$. Furthermore, let λ_f be an EV of $\mathbf{W}$. Then, $|\lambda_f| \leqslant C$. The proof is given in App. A.*

Theorem 4 deals with the EVs of $\mathbf{W}$. This has an implication to its corresponding SVs, which are used in (27). This implication stems from Theorem 1. Let us set $f = N$ in Eq. (8). Then,

$$\xi_N \geqslant |\lambda_N| \quad . \tag{29}$$

Now, following Theorem 4,
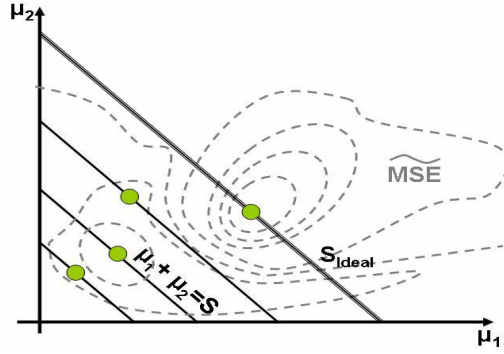
$$\lambda_N = C \quad . \tag{30}$$

Fig. 2. $\widetilde{\text{MSE}}$ as a function of $\{\mu_f\}_{f=1}^{N}$. Each green dot marks the vector $\vec{\mu}$ that minimizes $\widetilde{\text{MSE}}$, when S is fixed. The highlighted line marks the ideal value of S. The green dot along this line marks the global minimum of $\widetilde{\text{MSE}}$. This global minimum is derived in closed form and is thus unaffected by local minima.

Hence, using Eqs. (28,29,30)

$$\mu_N \triangleq \xi_N^2 \geqslant C^2 \quad . \tag{31}$$

Eq. (31) constrains the largest SV of $\mathbf{W}$. Therefore, we separate it from the rest of the SVs in Eq. (27). The advantage of this move will be clarified shortly. Thus we rewrite Eq. (27) as

$$\widetilde{\text{MSE}} = \frac{1}{N\mu_N} + \frac{1}{N}\sum_{f=1}^{N-1}\frac{1}{\mu_f} \quad . \tag{32}$$

3.2.2. Optimality of the singular values

In this paper we seek the tight lower bound on $\text{MSE}_{\hat{\mathbf{i}}}$. To achieve this, we have first expressed $\widetilde{\text{MSE}}$ by Eq. (32), in terms of the SVs of $\mathbf{W}$. Define the vector $\vec{\mu} = (\mu_1, \mu_2, \ldots, \mu_N)$. We now seek to minimize Eq. (32) as a function of $\vec{\mu}$. Apparently, $\widetilde{\text{MSE}}$ in Eq. (32) is reduced if we simply increase any single element of $\vec{\mu}$, μ_f, while the rest are constant. Can we do this arbitrarily? The answer is *no*. The reason is that Eqs. (24,25,26) bound the domain of $\mathbf{W}$, hence bounding the domain of its SVs. Therefore, any μ_f cannot be arbitrarily increased; the elements of $\vec{\mu}$ are mutually coupled. To express the coupling, we first use a normalization, by defining

$$S \triangleq \sum_{f=1}^{N} \mu_f \tag{33}$$

and setting it to be a constant. Later we alleviate the need for this arbitrary normalization (see Fig. 2).

Now, we show that under the constraint $S \equiv \text{Const}$,

$$\min_{\mu_N}\left\{\frac{1}{N\mu_N} + \frac{1}{N}\sum_{f=1}^{N-1}\frac{1}{\mu_f}\right\} = \frac{1}{NC^2} + \frac{1}{N}\sum_{f=1}^{N-1}\frac{1}{\mu_f} \quad . \tag{34}$$

To understand this observation consider Fig. 3. The minimization in Eq. (34) is only over the value of μ_N. Let μ_N be decreased by some small quantity $\Delta\mu_N$. Then, to conserve S in Eq. (33),
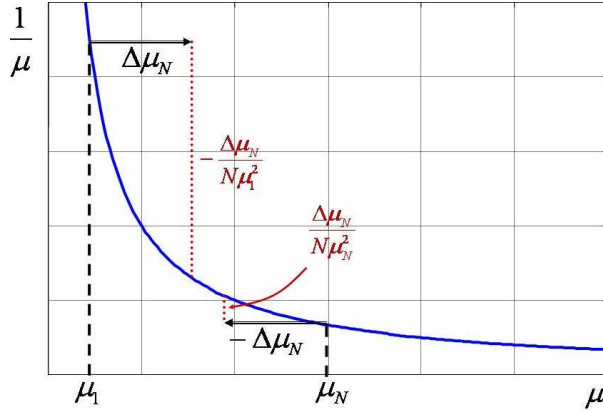
Fig. 3. The curve $1/\mu$ represents elements summed in Eq. (34). The black dashed lines mark a state of the SVs of $\mathbf{W}$. If the largest squared-SV, μ_N, is reduces by $\Delta\mu_N$ and in return the smallest squared-SV, μ_1, is increased by $\Delta\mu_N$, the sum in Eq. (34) is reduced.

the value of at least one other μ_f should increase. An increase of μ_f by $\Delta\mu_N$ reduces $\widetilde{\mathrm{MSE}}$. The reason is that the change induced by this increase is

$$\Delta\widetilde{\mathrm{MSE}} = \frac{\partial\widetilde{\mathrm{MSE}}}{\partial\mu_f}\Delta\mu_N = -\frac{1}{N\mu_f^2}\Delta\mu_N < 0 \quad . \tag{35}$$

We seek the strongest reduction of $\widetilde{\mathrm{MSE}}$. Clearly, in Eq. (35), the reduction $\Delta\widetilde{\mathrm{MSE}}$ is strongest if μ_f is the smallest element of the set $\{\mu_f\}_{f=1}^{N-1}$. The smallest element in this set is μ_1. Overall, decreasing μ_N by $\Delta\mu_N$ while increasing μ_1 by $\Delta\mu_N$ yields a net reduction of $\widetilde{\mathrm{MSE}}$ by

$$\Delta\widetilde{\mathrm{MSE}} = -\frac{\Delta\mu_N}{N}\left(\frac{1}{\mu_1^2} - \frac{1}{\mu_N^2}\right) < 0 \quad . \tag{36}$$

Thus, trading μ_N for μ_1 delivers a net benefit in $\widetilde{\mathrm{MSE}}$.

Since a benefit stems from a reduction of μ_N, then μ_N should be as low as possible. From Eq. (31), the lowest possible value of μ_N is C^2. To conclude,

Corollary 1. *In an optimal multiplexing code* $\mathbf{W}$*, the largest SV satisfies*

$$\xi_N = C \tag{37}$$

i.e.,

$$\mu_N = C^2 \quad . \tag{38}$$

This proves Eq. (34).

After determining μ_N, we turn to the other elements of $\vec{\mu}$. A trivial manipulation yields

$$\frac{1}{N}\sum_{f=1}^{N-1}\frac{1}{\mu_f} \equiv \frac{N-1}{N}\left(\frac{1}{N-1}\sum_{f=1}^{N-1}\frac{1}{\mu_f}\right) \quad . \tag{39}$$

The parentheses on the right hand side of (39) express the reciprocal harmonic mean of $\left\{\mu_f\right\}_{f=1}^{N-1}$. The *inequality of means* [34] states that:

$$\frac{1}{N-1}\sum_{f=1}^{N-1}\frac{1}{\mu_f} \geqslant \frac{N-1}{\Sigma_{f=1}^{N-1}\mu_f} \quad . \tag{40}$$

Using Eqs. (33,38) in Eq. (40) yields,

$$\frac{1}{N-1}\sum_{f=1}^{N-1}\frac{1}{\mu_f} \geqslant \frac{N-1}{S-C^2} \quad . \tag{41}$$

Combining Eqs. (34,39,41) into (32) yields

$$\widetilde{\text{MSE}} \geqslant B \quad , \tag{42}$$

while B is the lower bound of $\widetilde{\text{MSE}}$, given by

$$B = \frac{1}{NC^2} + \frac{(N-1)^2}{N(S-C^2)} \quad . \tag{43}$$

The lower bound B is constant as long as S is fixed. We wish that $\widetilde{\text{MSE}}$ will actually attain this fixed lower bound. Equality in Eq. (42) is obtained if equality holds in (40). This occurs when $\left\{\mu_f\right\}_{f=1}^{N-1}$ are all equal. Therefore, using Eq. (6), the elements of the set $\left\{\xi_f\right\}_{f=1}^{N-1}$ are also equal. We recap with the following corollary.

Corollary 2. *An ideal multiplexing matrix* **W** *is such, that all its SVs (but the largest one)* $\left\{\xi_f\right\}_{f=1}^{N-1}$ *are equal to each other. Its largest SV equals C.*

Note that we refer to matrices that attain the lower bound B as *ideal*. It is not guaranteed however, that such matrices exist for all values of N and C. Matrices that minimize $\widetilde{\text{MSE}}$ without reaching B are simply referred to as *optimal*.

3.2.3. The ideal variable S

We have shown in Sec. 3.2.2 that for a specific S, the MSE is minimized by matrices that comply with the terms in Corollary 2. We now relieve the constraint of an arbitrarily fixed S. Hence, S may vary (see Fig. 2). Furthermore, we derive the best value for S. This yields a condition on the elements of **W**. Following Lemma 1,

$$\text{trace}\left(\mathbf{W}^t\mathbf{W}\right) = \text{trace}\left(\mathbf{W}\mathbf{W}^t\right) \quad . \tag{44}$$

The diagonal elements of $\mathbf{W}\mathbf{W}^t$ are

$$\left(\mathbf{W}\mathbf{W}^t\right)_{m,m} = \sum_{s=1}^{N}\left(w_{m,s}\right)^2 \quad . \tag{45}$$

Due to Eq. (21),

$$\left(w_{m,s}\right)^2 \leqslant w_{m,s} \quad . \tag{46}$$

From Eqs. (20,45,46),

$$\left(\mathbf{W}\mathbf{W}^t\right)_{m,m} \leqslant \sum_{s=1}^{N}w_{m,s} = C \quad . \tag{47}$$

Following Lemma 2,

$$S = \sum_{k=1}^{N} \mu_k = \text{trace}\left(\mathbf{W}^t\mathbf{W}\right) \quad . \tag{48}$$

From Eqs. (44,47,48),

$$S = \text{trace}\left(\mathbf{W}\mathbf{W}^t\right) = \sum_{m=1}^{N} \left(\mathbf{W}\mathbf{W}^t\right)_{m,m} \leqslant NC \quad . \tag{49}$$

As indicated earlier in this section, S may vary. From Eq. (49) the domain of S is

$$0 \leq S \leq NC \quad . \tag{50}$$

Now it is possible to find the global optimum in this domain. From Eq. (43), B is minimized by using the largest S possible. Hence,

$$S_{\text{ideal}} = NC \quad . \tag{51}$$

Note that equality holds in (49) *if* and *only if* equality also holds in Eq. (46). Trivially, this happens if all elements of $\mathbf{W}$ are either 1's or 0's. In this case Eq. (49) yields Eq. (51), alleviating the need for an arbitrarily fixed S.

From Eqs. (43,49,51)

$$B \geqslant B_{\text{min}} \tag{52}$$

where

$$B_{\text{min}}(C) = \left[\frac{1}{NC^2} + \frac{(N-1)^2}{N(NC - C^2)}\right] \quad . \tag{53}$$

Equality holds if $w_{m,s} \in \{0,1\} \quad \forall m, s$. In other words, B reaches its lowest potential value $B_{\text{min}}(C)$, if S is given by Eq. (51).

Corollary 3. *The lower bound B_{min} of $\widetilde{\text{MSE}}$ is achieved using binary multiplexing matrices $\mathbf{W}$.*

Combining Corollary 3 and Eq. (20) results in

Theorem 5. $\widetilde{\text{MSE}}$ *yields B_{min} only if $C \in \mathbb{Z}^+$.*

We now combine the results of Corollaries 2 and 3. From Eqs. (3,42,43,52,53),

$$\text{MSE}_{\hat{\mathbf{i}}} \geqslant \sigma_a^2 B_{\text{min}}(C) \quad . \tag{54}$$

Theorem 6. *Equality in Eq. (54) is obtained by a matrix $\mathbf{W}$, if and only if this matrix complies with both Corollaries 2 and 3.*

The result in Eqs. (53,54) is very useful. Being a tight bound, B_{min} determines the behavior of $\text{MSE}_{\hat{\mathbf{i}}}$ as a function of both N and C. Furthermore, if $\mathbf{W}$ satisfies the optimality conditions, then $\sigma_a^2 B_{\text{min}}$ is exactly the expected value of $\text{MSE}_{\hat{\mathbf{i}}}$. In Fig. 4 we illustrate the behavior of B_{min} for a specific value of N and a range of values of C.
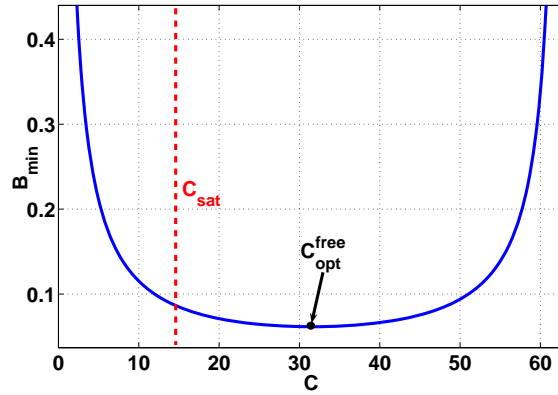
Fig. 4. The bound $B_{\min}(C)$, for $N = 63$. Here C varies from 1 to 63. The minimum of $B_{\min}(C)$ is at C_{opt}. There may exist C_{sat}, above which saturation occurs, inhibiting multiplexing.

4. The case of a free variable C

Sec. 3 dealt with power-regulated multiplexing, i.e., a fixed C. It introduced the lower bound on $\mathrm{MSE}_{\hat{\mathbf{i}}}$ for all pairs of $\{N, C\}$, via Eqs. (53,54). However, what happens if C is a free variable in the domain $C \in [1, N]$? For a given N, which value of $C \in [1, N]$ minimizes $\mathrm{MSE}_{\hat{\mathbf{i}}}$? This is discussed in this section.

The variable $\mathrm{MSE}_{\hat{\mathbf{i}}}$ is influenced by C through three mechanisms.

1. The parameter C affects the EVs of the optimal $\mathbf{W}$. This effect is reflected in the presence of C in the term for the bound $B_{\min}$ (see Eq. 53).

2. The parameter C may not be free to vary through all the domain $[1, N]$ due to saturation. There may exist C_{sat} such that $\forall\, C > C_{\mathrm{sat}}$, elements of the vector $\mathbf{a}$ (Eq. 1) become saturated. Such elements of $\mathbf{a}$ prevent the correct reconstruction of $\hat{\mathbf{i}}$ by Eq. (2).

3. The parameter C may affect σ_a in Eq. (3). So far in this paper, there was no consideration of variations in σ_a with respect to C. We shall see in Sec. 7 that this is often not the case, *e.g.* due to photon noise.

This section accounts only for mechanism 1, when discussing the optimal value of C. The other mechanisms are discussed in later sections. This allows the current section to compare our results to prior codes that do not consider mechanisms 2 and 3 as well, specifically, Hadamard codes [9, 18].

4.1. Minimization of $B_{\min}$

Recalling Sec. 3.1, if C is free to vary, then optimal matrices $\mathbf{W}$ are based on the Hadamard matrices and are known as the $\mathbf{S}$-matrices [9]. These matrices indeed comply with the conditions of Theorem 6. However, $\mathbf{S}$-matrices exist only for a subset of the possible values of N.

Define the domain

$$\Psi \triangleq \left\{ [1, N] \subset \mathbb{R}^+ \right\} \ . \tag{55}$$

In this domain, the value of the variable C that minimizes Eq. (53) is defined as

$$C_{\mathrm{opt}}^{\mathrm{free}} \triangleq \arg\min_{C \in \Psi} B_{\min}(C) \ . \tag{56}$$

Generally, $C_{\text{opt}}^{\text{free}}$ is not an integer. If it is not an integer, then it does not satisfy Theorem 5. Hence, this value does not correspond to a multiplexing matrix that meets the bound $B_{\min}(C_{\text{opt}}^{\text{free}})$. Nevertheless, $C_{\text{opt}}^{\text{free}}$ is important: it may imply an integer C (close to $C_{\text{opt}}^{\text{free}}$) that is associated with an ideal multiplexing matrix.

To find C_{opt}, let us null the derivative of $B_{\min}(C)$ (Eq. 53) with respect to C,

$$\frac{\partial B_{\min}}{\partial C} \equiv -\frac{2}{NC^3} - \frac{(N-1)^2(N-2C)}{N(NC-C^2)^2} = 0 \quad . \tag{57}$$

Nulling Eq. (57) yields

$$C_{\text{opt}}^{\text{free}} = \frac{N^2 - 2N - 3 \pm (N-1)\sqrt{(N^2 - 2N + 9)}}{4N - 8} \quad . \tag{58}$$

Apparently, Eq. (58) has two solutions. However, only one of them is in Ψ. Hence, $C_{\text{opt}}^{\text{free}}$ is unique. To satisfy Theorem 5, only integer C can be used. Let us define

$$C_{\text{opt}}^{\text{int}} \triangleq \text{ROUND}(C_{\text{opt}}^{\text{free}}) \tag{59}$$

as the integer value that is closest to C_{opt}. Define the integer domain

$$\Psi^{\text{int}} \triangleq \left\{ [1,N] \subset \mathbb{Z}^+ \right\} \quad . \tag{60}$$

We wish to know if $C_{\text{opt}}^{\text{int}}$ satisfies

$$C_{\text{opt}}^{\text{int}} = \arg \min_{C \in \Psi^{\text{int}}} B_{\min}(C) \quad . \tag{61}$$

This is discussed next.

4.2. Consistency with Hadamard matrices

Now we show that $C_{\text{opt}}^{\text{int}}$ degenerates to C_{Had}. It can be shown that

$$C_{\text{Had}} - C_{\text{opt}}^{\text{free}} = \frac{N-1}{N-2} \cdot \frac{N+1 - \sqrt{(N-1)^2 + 8}}{4}. \tag{62}$$

From Eq. (62), it can be shown that

$$\lim_{N \to \infty} (C_{\text{Had}} - C_{\text{opt}}^{\text{free}}) = 0.5 \,. \tag{63}$$

More generally, as shown in Fig. 5,

$$C_{\text{Had}} > C_{\text{opt}}^{\text{free}} > C_{\text{Had}} - 0.5 \quad . \tag{64}$$

Applying Eq. (59) in Eq. (64),

$$C_{\text{opt}}^{\text{int}} = C_{\text{Had}} \quad , \tag{65}$$

for values of N that correspond to $\mathbf{S}$ matrices. This is consistent with Ref. [9], which proves the optimality of Hadamard based-codes. Hence, Eq. (61) indeed holds in these cases.

Let us interpret our results:

- Hadamard-based codes are special cases, which satisfy our analysis. Their value of $\widetilde{\text{MSE}}$ attains $B_{\min}(C_{\text{Had}})$.
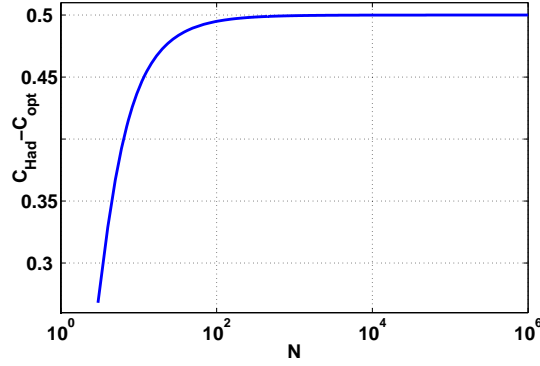- Eq. (61) is satisfied by Hadamard-based codes.

Fig. 5. A semi-logarithmic plot of Eq. (62).

- Our analysis generalizes theories that were previously developed for multiplexing. In particular, the analysis applies to cases (values of $\{N,C\}$) for which Hadamard-based codes do not exist.

We wish to clarify that *optimal* multiplexing codes do not necessitate an integer C. When the condition specified in Theorem 5 is not met, then $B \neq B_{\min}$. Thus, for a non-integer C, no matrix $\mathbf{W}$ is *ideal*, but a certain $\mathbf{W}$ can still be *optimal*, *i.e.* minimizing $\widetilde{\text{MSE}}$ within constraints (24,25,26).

5. Saturation

From Sec. 4, recall mechanism 2 by which C influences $\text{MSE}_{\hat{\imath}}$. It is saturation. Due to saturation, C cannot exceed C_{sat}. In this section we discuss seeking an optimal value of C, under the saturation constraint.

The unsaturated domain of C is

$$\Psi^{\text{sat}} = \{\Psi \cap [1, C_{\text{sat}}]\} \quad , \tag{66}$$

where Ψ is defined in Eq. (55). We seek the value

$$C_{\text{opt}}^{\text{sat}} = \arg \min_{C \in \Psi^{\text{sat}}} B_{\min}(C) \quad . \tag{67}$$

Let us discuss two cases. First, suppose that $C_{\text{opt}}^{\text{free}} \leq C_{\text{sat}}$, where $C_{\text{opt}}^{\text{free}}$ is defined in Eq. (58). In this case, $C_{\text{opt}}^{\text{free}} \in \Psi^{\text{sat}}$. Since $C_{\text{opt}}^{\text{free}}$ uniquely minimizes Eq. (53) in this domain, the optimal solution degenerates to that described in Sec. 4.1.

In the second case, $C_{\text{opt}}^{\text{free}} > C_{\text{sat}}$, i.e, $C_{\text{opt}}^{\text{free}} \notin \Psi^{\text{sat}}$. Then, there is no value of $C \in \Psi^{\text{sat}}$ at which the derivative of $B_{\min}(C)$ is null. Hence, the global minimum of the function $B_{\min}(C)$ is obtained at the constraint C_{sat}. This is illustrated, for example, in Fig. 4. To conclude,

$$C_{\text{opt}}^{\text{sat}} = \begin{cases} C_{\text{opt}}^{\text{free}} & \text{if } C_{\text{opt}}^{\text{free}} \leq C_{\text{sat}} \\ C_{\text{sat}} & \text{otherwise} \end{cases} \quad . \tag{68}$$

6. Some ideal solutions

Sec. 3 derived a lower bound on $\widetilde{\text{MSE}}$, which can be obtained by multiplexing. Sec. 3 also shows that in order to attain this bound, certain conditions should be satisfied by the multiplexing code.

However, it is not obvious that such *ideal* codes exist. The familiar Hadamard-based **S** matrices are ideal codes. Their value of $\widetilde{\text{MSE}}$ attains $B_{\min}(C_{\text{Had}})$. Are there other matrices, that satisfy these optimality conditions, for other values of $\{N,C\}$? In the following, we show that indeed there is a class of such matrices.

6.1. Strongly regular graphs as a solution

The adjacency matrix Ω described in Theorem 3 can be used as the desired multiplex matrix **W**, sought in Sec. 3, as we now detail:

- Since Ω is an adjacency matrix, it is binary. Hence, it satisfies Corollary 3.

- The highest SV of the desired **W** is C, with multiplicity which is generally 1 (See Corollary 2). Similarly, the highest SV of Ω is k, with the same multiplicity. Hence, we set $k = C$.

- Setting the parameters $\alpha = \beta$ in Eqs. (14,15) yields $|\lambda_1^\Omega| = |\lambda_2^\Omega| = \frac{\sqrt{\Delta}}{2}$. Using Eq. (18), all SVs of Ω are equal to $\frac{\sqrt{\Delta}}{2}$, except for the largest SV, which equals k. Namely,

$$\xi_1^\Omega = \sqrt{C - \beta} \tag{69a}$$

$$\xi_2^\Omega = C \quad . \tag{69b}$$

Therefore, Ω satisfies the conditions of Corollary 2. To summarize,

Theorem 7. *Let Ω be the adjacency matrix of a $(N;C;\alpha;\alpha)$ strongly regular graph. Then, Ω is an ideal matrix **W** solving Eqs. (23,24,25,26).*

Recall that an SRG is subject to a constraint given in Eq. (13) on the values of its parameters. If $\alpha = \beta$ as in Theorem 7, then Eq. (13) takes the following form

$$\alpha = \frac{k(k-1)}{(N-1)} = \frac{C(C-1)}{(N-1)} \quad . \tag{70}$$

As an example, consider the $(45;12;3;3)$ SRG, described in Refs. [35–37].This is a class of graphs, determined up to an isomorphism. The adjacency matrix of a representative of this class is shown as a binary image in Fig. 6. Note that the parameters of this graph satisfy Eq. (70). Additional examples for SRGs can be found in [36, 37].

6.2. Solutions from complement graphs

In Sec. 6.1 we have devised a set of solutions for the sought optimal **W**. Such a set may be derived indirectly from another set of SRGs. Refs. [31, 35] formalize the following theorem:

Theorem 8. *A graph G is strongly regular with parameters $(N;k;\alpha;\beta)$ if and only if its complement $\bar{G}$ is also strongly regular, with parameters $(N; N-k-1; N-2k+\beta-2; N-2k+\alpha)$.*

Theorem 8 means that given an SRG, $\bar{G}$ with some parameters $(N;\bar{C};\bar{\alpha};\bar{\alpha}+2)$, there exists a complement SRG, G, with parameters $(N;C;\alpha;\alpha)$. Here $C = N - \bar{C} - 1$ and $\alpha = N - 2\bar{C} + \bar{\alpha}$. Hence, using Theorem 7, G is a solution for Eqs. (23,24,25,26).

6.3. A simulated example

To illustrate the use of an SRG in multiplexing, we describe a simulation of spectrometry. The vector **i** represents noiseless graylevel measurements, where each element of the array
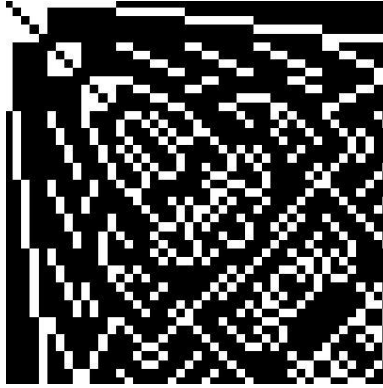
Fig. 6. An example for an adjacency matrix of an SRG with parameters $(45; 12; 3; 3)$ developed by Ref. [35]. Here 1s and 0s are represented by white and black squares, respectively. Here $\alpha = \beta = 3$, hence this graph satisfies the conditions of Theorem 7.

expresses the radiance in a narrow spectral band, as transmitted through the atmosphere. As "ground-truth" $\mathbf{i}$, we took the data in [38]. These graylevel values are plotted in Fig. 7a.

We simulated noisy infrared measurements of these values. They were compounded by Gaussian white noise, having $\sigma_a = 8$ graylevels independent of the signal, and quantized by an 8-bit camera, having a range of [0,255] graylevels. The resulting absolute error $|\hat{i} - i|$ for this trivial measurement process is plotted by a red line in Fig. 7b.

Then, we simulated spectral multiplexing based on an SRG. In this data, $N = 45$. Thus, we used the adjacency matrix shown in Fig. 6 for multiplexing the spectral bands. Thus, in each measurement, 12 narrow spectral bands were simultaneously sensed. Then, noise and quantization were applied as in the standard acquisition process. There was no saturation in $\mathbf{a}$: considering the values $\mathbf{i}$ plotted in Fig. 7a, all elements of $\mathbf{a}$ were lower than 255 graylevels. The measurements were then demultiplexed, yielding a less noisy estimate $\hat{\mathbf{i}}$. The resulting absolute error $|\hat{i} - i|$ of the demultiplexed values is plotted by a green line in Fig. 7b. Indeed, the noise is significantly reduced. Quantitatively, the noise reduction is as expected: empirically, $\mathrm{MSE}_{\hat{\mathbf{i}}}$ is 6.96 squared-graylevels, when averaged over many such randomly noised simulations. This is consistent with the expected theoretical value of $\mathrm{MSE}_{\hat{\mathbf{i}}}$, obtained by Eqs. (53,54).

7. Photon noise

In this section, we discuss mechanism 3 by which C influences $\mathrm{MSE}_{\hat{\mathbf{i}}}$, as mentioned in Sec. 4. Here we deal with the presence of photon noise in the acquired measurements. Generally, photon noise inhibits the multiplexing gain. The variance of photon noise increases linearly with the acquired signal. Hence, an increase of radiance used in each measurement increases the noise. This effect degrades the multiplexing gain, sometimes even causing it to be counterproductive [22, 24].

7.1. The affine noise model

As in [22], we use the *affine noise model*. It exists in high grade detectors, which have a linear radiometric response. The noise can be divided into two components, signal-dependent and signal-independent. Regardless of the photon flux, signal-independent noise is created by dark current [24, 39, 40], amplifier noise and the quantizer in the sensor circuity [40]. Denote the gray-level variance of the signal-independent noise by κ_{gray}^2.
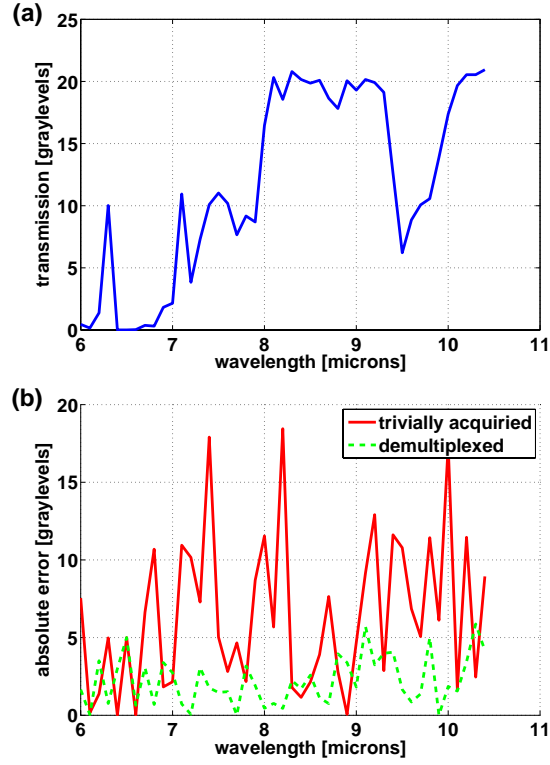
Fig. 7. (a) Graylevel spectral radiance of light transmitted through the atmosphere [38]. It is used to simulate the ground truth **i**. (b) Absolute error values $|\hat{i} - i|$ of simulated estimates based either on trivial sensing (red plot) or on SRG-based multiplexing (green plot).

Fundamental signal-dependent noise is related to two random effects. The photon flux and the uncertainty of the photon-electron conversion process which occurs in the detector. Overall, the random number $n_{\mathrm{electr}}^{\mathrm{photo}}$ of photo-generated electrons is Poisson distributed [39, 41, 42]. In this distribution, the variance of $n_{\mathrm{electr}}^{\mathrm{photo}}$ is

$$\mathrm{VAR}(n_{\mathrm{electr}}^{\mathrm{photo}}) = \mathscr{E}(n_{\mathrm{electr}}^{\mathrm{photo}}) \quad , \tag{71}$$

where $\mathscr{E}$ denotes expectation. This variance linearly increases with the measured electric signal $n_{\mathrm{electr}}^{\mathrm{photo}}$. The number of detected electrons $n_{\mathrm{electr}}^{\mathrm{photo}}$ is proportional to the gray-level of the acquired pixel value a

$$a = n_{\mathrm{electr}}^{\mathrm{photo}}/Q_{\mathrm{electr}} \quad . \tag{72}$$

Here Q_{electr} is the number of photo-generated electrons required to change a unit gray-level. Typically $Q_{\mathrm{electr}} \gg 1$. Combining Eqs. (71,72) yields a variance in gray levels

$$\mathscr{E}(n_{\mathrm{electr}}^{\mathrm{photo}})/Q_{\mathrm{electr}}^2 = a/Q_{\mathrm{electr}} \quad . \tag{73}$$

Compounded with signal-independent noise, the total noise variance of the measured gray-level [24, 39] is

$$\sigma_a^2 = \kappa_{\mathrm{gray}}^2 + a/Q_{\mathrm{electr}} \quad . \tag{74}$$

Now, consider a scenario in which each radiation source, s yields a similar image radiance, i_s. Following Eqs. (1) and (20), in each measurement the acquired value is

$$a \approx i_s C \ . \tag{75}$$

Thus Eq. (74) can be rephrased as

$$\sigma_a^2 = \kappa_{\text{gray}}^2 + C\eta^2 \quad . \tag{76}$$

Here $\eta^2 \approx i_s/Q_{\text{electr}}$ is the photon noise variance, induced by i_s. Eq. (76) is an affine function of the number of active sources C. This was demonstrated experimentally in Ref. [22]. The parameters κ_{gray}^2 and η^2 depend on the setup hardware. A way to calibrate them is described in Ref. [22].

7.2. Optimal multiplexing

In this section we describe how to derive an optimal value of C, termed $C_{\text{opt}}^{\text{final}}$, accounting for all the effects. These include the various noise mechanisms and saturation. In addition, the section describes an approach to obtain a multiplexing matrix $\mathbf{W}$ that corresponds to $C_{\text{opt}}^{\text{final}}$.

From Eqs. (3,22),

$$\text{MSE}_{\hat{\mathbf{i}}} = \sigma_a^2 \widetilde{\text{MSE}} \ . \tag{77}$$

Embedding the noise variance of Eq. (76) into Eq. (77) yields

$$\text{MSE}_{\hat{\mathbf{i}}} = (\kappa_{\text{gray}}^2 + C\eta^2) \widetilde{\text{MSE}} \ . \tag{78}$$

Ref. [22] suggested a minimization of the $\text{MSE}_{\hat{\mathbf{i}}}$ expression given in Eq. (78). It consists of the following steps.

1. Scan the range of C values from 1 to $C_{\text{opt}}^{\text{sat}}$, where $C_{\text{opt}}^{\text{sat}}$ is defined in Eq. (68). For each value of C, perform the subsequent steps 2 and 3.

2. Find the matrix $\mathbf{W}(C)$ that optimizes $\widetilde{\text{MSE}}$, defined in Eq. (22). This optimization is constrained by (24,25,26).

3. Based on $\mathbf{W}(C)$, calculate the expected multiplex gain $\text{MSE}_{\hat{\mathbf{i}}}(C)$ using Eqs. (22,78).

4. Let

$$C_{\text{opt}}^{\text{final}} = \arg\min_{C} \ \text{MSE}_{\hat{\mathbf{i}}}(C) \ . \tag{79}$$

5. The desired multiplexing code is $\mathbf{W}(C_{\text{opt}}^{\text{final}})$.

Note that the result of this process is a truly optimal multiplexing matrix, accounting for both saturation and photon-noise. Furthermore, in step 1, there is no necessity for exhaustive search of $\text{MSE}_{\hat{\mathbf{i}}}(C)$ for all $C \in [1, C_{\text{opt}}^{\text{sat}}]$. The reason is that $\text{MSE}_{\hat{\mathbf{i}}}(C)$ is well behaved [22], hence one can incorporate efficient optimization procedures. The core of the process is step 2. In Ref. [22], this minimization was based on a numerical search of $\widetilde{\text{MSE}}$, i.e., of Eqs. (23,24,25,26). However, Eq. (22) is not unimodal. Hence, it is difficult to guarantee that a global minimum of $\widetilde{\text{MSE}}$ is numerically reached.

Following Secs. 3 and 6 above, we may simplify step 2 in two ways. The first way is to create ideal multiplexing codes $\mathbf{W}(C)$ using adjacency matrices of known SRGs having the proper parameters (see Sec. 6). In this way, step 2 avoids numerical search altogether. However, SRGs are not available for every $\{N,C\}$ set. Thus, numerical optimization may still be needed.

Nevertheless, there is a second, more general way by which step 2 is simplified: $B_{\min}(C)$ can set the termination of a numerical search, as is explained in the following.

Any numerical optimization process is iterative. Hence, let the above mentioned step 2 encapsulate an iterative process. At the l'th iteration, there is a matrix $\mathbf{W}_l$ that complies with Eqs. (24,25,26). Based on Eq. (22), a corresponding value $\widetilde{\mathrm{MSE}}_l$ is derived. At iteration $l+1$, the multiplexing matrix changes to $\mathbf{W}_{l+1}$, with a corresponding $\widetilde{\mathrm{MSE}}_{l+1}$. The numerical optimization seeks, in general, to yield $\widetilde{\mathrm{MSE}}_{l+1} < \widetilde{\mathrm{MSE}}_l$, hence minimizing $\widetilde{\mathrm{MSE}}$ as the iterations progress. A numerical minimization of $\widetilde{\mathrm{MSE}}$ can be efficient using a gradient-based method [22]. This is done since the gradient of $\widetilde{\mathrm{MSE}}$ (with respect to the elements of $\mathbf{W}_l$) is given in closed-form (See [22]).

In iterative optimization, an important matter is to know when to terminate it, i.e., at which l does $\mathbf{W}_l$ yield $\widetilde{\mathrm{MSE}}_l$ that is indeed close enough to its true optimum? Here the conditions and bounds we obtained in Sec. 3.2 come into play. Specifically, $B_{\min}(C)$ can terminate the iterations. If at the l'th iteration $\widetilde{\mathrm{MSE}}_l \approx B_{\min}(C)$, then the iterations may stop: $\mathbf{W}_l$ is as good as it can get. Hence, it can be set as $\mathbf{W}(C)$ in step 2 above.

8. Discussion

We derived several lower bounds on the estimation error, in conjunction with conditions required to attain them. The bound B given in Eq. (43) determines best feasible performance by any code for which the sum of SVs is arbitrarily fixed. Correspondingly, Corollary 2 states the structure of desired codes. A stronger bound is $B_{\min}(C)$, given by Eq. (53), as it states the best conceivable performance of any code, whatever its SVs are. However, finding a code that attains $B_{\min}(C)$ is more difficult, since such a code is constrained to satisfy Eqs. (48,51). Such a constraint can only be satisfied by binary codes (Corollary 3). This lead to an interesting link to graph theory, which is usually a field distinct from analysis of intensity noise. We showed that SRG have adjacency matrices that are ideal, in the sense that they attain $B_{\min}(C)$.

An important contribution is the algorithm described in Sec. 7.2. It exploits the performance bounds derived in the paper for a practical task: validating and influencing a numerical search for optimal multiplexing codes. These codes can have any number of sources N, activated sources C, and be tailored to the characteristics of noise and saturation.

The analysis of general matrices $\mathbf{W}$ is helpful to subcases of such matrices, be they binary [3,4,9–12,26] and/or cyclic [8, 9, 20, 23]. Furthermore, the analysis may be extended to noise models [24] that are different than Eq. (76).

We believe that these results are important to the wide range of fields where the multiplexing principle is used. The results can help yield measurement devices for estimating an arbitrary number of sought variables, where the measurement apparatus is optimal under fundamental and practical limitations and effects.

A. Proof of Theorem 4

To make the paper self-contained, we prove Theorem 4. This derivation is a variation of a proof that appears in [33]. Define

$$\mathbf{W}^{\mathrm{offset}} \triangleq \mathbf{W} - C\mathbf{I} \ . \tag{80}$$

Let $w_{m,s}^{\mathrm{offset}}$ be the elements of $\mathbf{W}^{\mathrm{offset}}$. Following Eq. (20),

$$\sum_{s=1}^{N} w_{m,s}^{\mathrm{offset}} = 0 \quad \forall m \in \{1,2,\ldots,N\} \ . \tag{81}$$

Hence, one of the EVs of $\mathbf{W}^{\text{offset}}$ is 0. In other words, $\det(\mathbf{W}^{\text{offset}}) = 0$. By the definition in (80), this means that $\det(\mathbf{W} - C\mathbf{I}) = 0$, *i.e.*, C is an EV of $\mathbf{W}$. This proves the first part of Theorem 4.

Suppose that $\mathbf{u}^f = (u_1^f, \ldots, u_N^f)^t$ is an eigenvector corresponding to λ_f. Without the loss of generality, we may normalize $\mathbf{u}^f$, such that

$$\max_{m \in \{1,\ldots,N\}} |u_m^f| = 1 \ . \tag{82}$$

Hence,

$$u_m^f = 1 \tag{83}$$

for a certain $m \in \{1,\ldots,N\}$. From Eqs. (21,82)

$$|w_{m,s}u_m^f| \leqslant 1 \qquad \forall m, s \in \{1,\ldots,N\} \ . \tag{84}$$

Eqs. (20,84) directly lead to a constraint on the absolute value of the *s*'th component of $\mathbf{W}\mathbf{u}^f$.

$$|(\mathbf{W}\mathbf{u}^f)_m| = |\sum_{s=1}^{N} w_{m,s}u_s^f| \leqslant \sum_{s=1}^{N} |w_{m,s}u_s^f| \leqslant C \ . \tag{85}$$

In addition, $\mathbf{u}^f$ is an eigenvector of $\mathbf{W}$. Hence,

$$|(\mathbf{W}\mathbf{u}^f)_m| = |\lambda_f \mathbf{u}^f|_m = |\lambda_f u_m^f| \ . \tag{86}$$

From Eq. (83), Eq. (86) becomes

$$|(\mathbf{W}\mathbf{u}^f)_m| = |\lambda_f| \ . \tag{87}$$

Using Eq. (87) in (85) yields

$$|\lambda_f| \leqslant C \ . \tag{88}$$

This proves the second half of Theorem 4.

Acknowledgments