

2016-12-01

Performance analysis of low-flux least-squares single-pixel imaging

Donggeek Shin, Jeffrey H Shapiro, Vivek K Goyal. 2016. "Performance Analysis of Low-Flux Least-Squares Single-Pixel Imaging." IEEE SIGNAL PROCESSING LETTERS, Volume 23, Issue 12, pp. 1756 - 1760. <https://doi.org/10.1109/LSP.2016.2617329>

<https://hdl.handle.net/2144/39192>

"Downloaded from OpenBU. Boston University's institutional repository."

Performance Analysis of Low-Flux Least-Squares Single-Pixel Imaging

Donggeek Shin, *Student Member, IEEE*, Jeffrey H. Shapiro, *Life Fellow, IEEE*, and Vivek K Goyal, *Fellow, IEEE*

Abstract—A single-pixel camera is able to computationally form spatially-resolved images using one photodetector and a spatial light modulator. The images it produces in low light-level operation are imperfect, even when the number of measurements exceeds the number of pixels, because its photodetection measurements are corrupted by Poisson noise. Conventional performance analysis for single-pixel imaging generates estimates of mean-squared error from Monte Carlo simulations, which require long computational times. In this letter, we use random matrix theory to develop a closed-form approximation to the mean-squared error of the widely used least-squares inversion method for Poisson noise-limited single-pixel imaging. We present numerical experiments that validate our approximation and a motivating example showing how our framework can be used to answer practical optical design questions for a single-pixel camera.

Index Terms—Single-pixel imaging, Poisson noise, least-squares estimator, performance analysis, random matrix theory, Marchenko–Pastur law.

I. INTRODUCTION

A conventional digital camera forms an n -pixel image by using an array of n photodetectors to directly and separately measure the light falling on each image-plane pixel. A single-pixel camera, on the other hand, masks the light reaching an image-plane-covering photodetector with a sequence of m spatial patterns—produced by an n -pixel spatial light modulator (SLM), such as a digital micromirror device—and then solves the resulting inverse problem to produce its n -pixel image [1]. When detector arrays are very costly, single-pixel cameras can offer inexpensive alternatives to conventional cameras. In addition, they may be the *only* viable imaging approach at optical wavelengths for which high-resolution array detectors are unavailable [2].

At sufficiently low light levels, a high-sensitivity photodetector records discrete photon counts, whose Poisson statistics characterize its dominant noise source [3]. High-quality imaging at low light levels is important for applications such as range imaging [4]–[8] and photosensitive cell imaging [9]. Whereas some of the literature on low-light imaging with single-pixel cameras has employed compressed-sensing estimators, in which images are formed from $m < n$ measurements [10], [11], better low light-level performance with such cameras is obtained from $m \geq n$ measurements [12].

This material is based upon work supported in part by a Samsung Scholarship, and by the US National Science Foundation under Grant No. 1422034.

D. Shin and J. H. Shapiro are with the Department of Electrical Engineering and Computer Science and the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

V. K. Goyal is with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA.

Computing the mean-squared error (MSE) of even the simplest estimators for low light-level single-pixel imaging is non-trivial, owing to the difficulties introduced by the Poisson statistics of photon counting and the non-Gaussian nature of the camera’s SLM patterns. Consequently, their MSEs are conventionally estimated using Monte Carlo simulations [13], despite the attendant high computational cost.

Recently, random matrix theory (RMT) has been used for accurate, closed-form characterization of estimation errors in many contexts, including channel identification [14], wireless communications [15], and inverse covariance estimation [16]. Two novel constraints prevent direct application of existing RMT results to low light-level single-pixel imaging using random SLM patterns:

- 1) Low light-level single-pixel imaging measurements have signal-dependent Poisson distributions. This observation model deviates dramatically from the noiseless or additive, signal-independent Gaussian noise models employed in prior work.
- 2) Prior error analyses presume zero-mean observations that allow direct application of classical RMT results. The Poisson-distributed observations in low light-level single-pixel imaging, however, have non-zero signal-dependent mean values, because the SLM-patterned light intensities being detected are non-negative.

In this letter, we derive a closed-form approximation for the MSE of least-squares single-pixel imaging that accounts for Poisson-distributed observations obtained with the commonly-employed Bernoulli (randomly 0 or 1) SLM patterns. We borrow techniques from RMT to arrive at a simple expression that avoids the implicit characterizations obtained through the replica method [17], [18] and issues arising from the SLM patterns’ non-zero mean values. Our numerical experiments show that our approximation is nearly exact, with much lower computational cost than Monte Carlo methods and higher accuracy than a baseline asymptotic MSE approximation.

II. SINGLE-PIXEL IMAGING FRAMEWORK

Figure 1 illustrates our single-pixel imaging setup. We use $\mathbf{x} \in \mathbb{R}_+^{n \times 1}$ to denote the light-intensity vector that we aim to recover, e.g., the vectorized form of a two-dimensional scene. An n -pixel SLM images light coming from the scene onto a sequence of $m > n$ independent, identically-distributed (i.i.d.) random subsets of the pixels covering the camera’s photodetector. We use \mathbf{a}_k to denote the $n \times 1$ random vector for the k th SLM pattern, and we take its entries to be i.i.d. Bernoulli random variables with success probability p . By

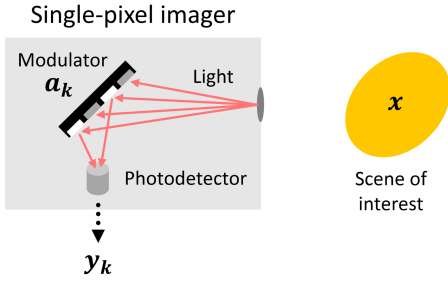


Fig. 1. Single-pixel imaging setup. For the k th measurement, the light intensity \mathbf{x} is spatially modulated, using the binary pattern \mathbf{a}_k , such that the light incident on the single photodetector has intensity $\mathbf{a}_k^T \mathbf{x}$. The photodetector's output then satisfies $\mathbf{y}_k \sim \text{Poisson}(\mathbf{a}_k^T \mathbf{x})$, and the process is repeated for m measurements. This figure depicts the spatial-light modulator operating in reflection, as is the case for a programmable micromirror device, but it could also work in transmission, as occurs with a random diffraction grating [19].

using low-flux photodetection statistics, we can write our k th photon-count observation as

$$\mathbf{y}_k \sim \text{Poisson}(\mathbf{a}_k^T \mathbf{x}), \quad k = 1, 2, \dots, m, \quad (1)$$

meaning that it has the probability mass function

$$\Pr(\mathbf{y}_k | \mathbf{a}_k, \mathbf{x}) = \frac{\exp\{-\mathbf{a}_k^T \mathbf{x}\} (\mathbf{a}_k^T \mathbf{x})^{\mathbf{y}_k}}{\mathbf{y}_k!}.$$

We write this more compactly as

$$\mathbf{y} \sim \text{Poisson}(\mathbf{A}\mathbf{x}), \quad (2)$$

where \mathbf{y} is the $m \times 1$ observation vector, \mathbf{A} is the $m \times n$ random matrix that concatenates the $1 \times n$ row vectors $\mathbf{a}_1^T, \dots, \mathbf{a}_m^T$, and $\text{Poisson}(\cdot)$ is applied entrywise.

Single-pixel imaging's inverse problem is to reliably estimate the size- n intensity \mathbf{x} from the observations \mathbf{y} and knowledge of the SLM pattern matrix \mathbf{A} . Were $m = n$ and \mathbf{A} the $n \times n$ identity matrix \mathbf{I}_n , (2) would simply describe a pixel-by-pixel scanning imager. Alternatively, the case of $m = n$ with $\mathbf{A} \neq \mathbf{I}_n$ but nonsingular has been termed multiplexed imaging [20]. Its analysis has supported the use of Hadamard matrices [21], [22], and it has been shown that constraining \mathbf{x} to be non-negative leads to a multiplexing advantage in low-flux operation [23].

Because this letter addresses low-flux operation, we will assume the condition $m > n$ that is effective in noise mitigation. Our inverse problem is thus classical, rather than compressive, where a prior on \mathbf{x} is required to compute a solution. We study the least-squares (pseudo-inverse) estimator,

$$\hat{\mathbf{x}}_{\text{LS}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}, \quad (3)$$

where we assume that $\text{rank}(\mathbf{A}) = n$, so that $\mathbf{A}^T \mathbf{A}$ is invertible. The least-squares method is not maximum likelihood; for example, it fails to enforce light intensity's non-negativity. Nevertheless, despite its model mismatch, the least-squares estimator is a popular inversion method because of its low computational complexity and its exactness as $m \rightarrow \infty$. Moreover, the often-used correlation estimator, also known as the ghost imaging estimator, for single-pixel imaging from overcomplete photon-count observations is a variant of the least-squares estimator [24].

The least-squares imager's MSE,

$$\text{mse}(\mathbf{x}, \hat{\mathbf{x}}_{\text{LS}}) = \mathbb{E}[\|\mathbf{x} - \hat{\mathbf{x}}_{\text{LS}}\|_2^2],$$

can be approximated from Monte Carlo simulations, but highly-accurate results require long simulation runs (many Monte Carlo trials). Thus, our goal is an accurate closed-form MSE approximation that is cheap in computation.

III. APPROXIMATING THE MSE

A. Derivation

Our first step is to express the MSE in terms of \mathbf{A} and \mathbf{x} . Given \mathbf{A} and \mathbf{x} , the observation \mathbf{y} has statistically-independent, Poisson-distributed components, so that its mean vector is $\mathbf{A}\mathbf{x}$, and its covariance matrix is diagonal with $\mathbf{A}\mathbf{x}$ being the $m \times 1$ vector of its non-zero elements. It follows that we can write $\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a zero-mean random vector with covariance matrix $\text{diag}(\mathbf{A}\mathbf{x})$. In the Appendix we show that the MSE can be expressed as

$$\text{mse} = \text{Tr}(\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^T \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T]), \quad (4)$$

where $\text{Tr}(\cdot)$ denotes the trace of a square-matrix argument.

Using cmse to denote the conditional MSE given the random-pattern matrix \mathbf{A} now gives us

$$\begin{aligned} \text{cmse} &= \text{Tr}(\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^T | \mathbf{A}] \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T) \\ &= \text{Tr}(\text{diag}(\mathbf{A}\mathbf{x}) \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T). \end{aligned} \quad (5)$$

Exact computation of the expectation of cmse over \mathbf{A} is complicated by the statistical dependence of $\text{diag}(\mathbf{A}\mathbf{x})$ and $\mathbf{A}(\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T$. Approximating them as independent terms, for large problem sizes, leads to

$$\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^T \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T] \approx \mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^T] \mathbb{E}[\mathbf{A}(\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T]. \quad (6)$$

We then observe that

$$\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^T] = \mathbb{E}[\text{diag}(\mathbf{A}\mathbf{x})] = p\|\mathbf{x}\|_1 \mathbf{I}_m, \quad (7)$$

where the first equality uses iterated expectation conditioned on \mathbf{A} , and the second equality uses the Bernoulli distribution of \mathbf{A} 's entries. Thus, we can write

$$\begin{aligned} \text{mse} &= \text{Tr}(\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^T \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T]) \\ &\stackrel{(a)}{\approx} \text{Tr}(\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^T] \mathbb{E}[\mathbf{A}(\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T]) \\ &\stackrel{(b)}{=} p\|\mathbf{x}\|_1 \text{Tr}(\mathbb{E}[\mathbf{A}(\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T]), \\ &\stackrel{(c)}{=} p\|\mathbf{x}\|_1 \text{Tr}(\mathbb{E}[(\mathbf{A}^T \mathbf{A})^{-1}]), \end{aligned} \quad (8)$$

where (a) uses the independence approximation (6); (b) uses (7); and (c) uses the trace operation's being cyclic plus the fact that trace commutes with expectation.

At this point our MSE approximation for the least-squares estimator is the product of imaging parameters and the trace-expectation of the inverse of a scalar multiple of the correlation matrix $\mathbf{C} = \mathbf{A}^T \mathbf{A}/m$. Because the trace of a matrix equals the sum of its eigenvalues, our next task is to characterize the inverse correlation matrix's eigenvalues $\lambda(\mathbf{C}^{-1})$, so that we can compute the trace-expectation factor in $\text{mse} = p\|\mathbf{x}\|_1 \text{Tr}(\mathbb{E}[(\mathbf{A}^T \mathbf{A})^{-1}])$.

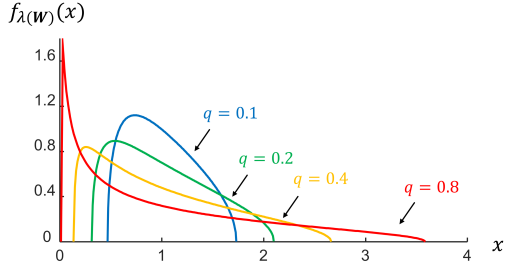


Fig. 2. Plots of Marchenko-Pastur law EDFs for different q values.

RMT's celebrated Marchenko-Pastur law [25] characterizes the eigenvalue density function (EDF) of the Wishart matrix, $\mathbf{W} = \mathbf{X}^T \mathbf{X}/m$ associated with an $m \times n$ random matrix \mathbf{X} with known parameterization.

(Marchenko–Pastur Law) Let \mathbf{X} be an $m \times n$ matrix whose entries are i.i.d. random variables with mean 0 and variance σ^2 . As $m, n \rightarrow \infty$ with $q = n/m \in (0, 1]$ fixed, the EDF $f_{\lambda(\mathbf{W})}$ of the Wishart matrix $\mathbf{W} = \mathbf{X}^T \mathbf{X}/m$ converges to

$$f_{\lambda(\mathbf{W})}(x) = \frac{1}{2\pi\sigma^2} \frac{\sqrt{(q_+ - x)(x - q_-)}}{qx}, \quad (9)$$

where $q_{\pm} = \sigma^2(1 \pm \sqrt{q})^2$ and $x \in [q_-, q_+]$.

Figure 2 plots Marchenko-Pastur EDFs for several different q values. Using

$$\mu_{\lambda(\mathbf{W})}(k; m, n) = \int_{\text{supp}(f)} x^k f_{\lambda(\mathbf{W})}(x) dx \quad (10)$$

to denote the k th-order moment of the Wishart matrix's eigenvalues, we have from [15] that

$$\text{Tr}(\mathbb{E}[\mathbf{W}^k]) = n \mu_{\lambda(\mathbf{W})}(k; m, n). \quad (11)$$

Unfortunately, the Marchenko-Pastur law is not directly applicable to evaluating $\text{Tr}(\mathbb{E}[(\mathbf{A}^T \mathbf{A})^{-1}])$ because \mathbf{A} 's entries have non-zero means. To deal with this difficulty, we study variations on that law when the correlation matrix is generated from non-negative random vectors.

Let $\mathbf{A}' = \mathbf{A} - \mathbb{E}[\mathbf{A}]$, where $\mathbb{E}[\mathbf{A}]$ is the $m \times n$ matrix all of whose entries equal p , so that \mathbf{A}' comprises i.i.d. zero-mean random variables. Next, consider the Wishart matrix $\mathbf{C}' = (\mathbf{A}')^T \mathbf{A}'/m$. Then, because the single non-zero eigenvalue of $\mathbb{E}[\mathbf{A}]^T \mathbb{E}[\mathbf{A}]/m$, viz., $\lambda(\mathbb{E}[\mathbf{A}]^T \mathbb{E}[\mathbf{A}]/m) = np^2$, is much larger than $\lambda_{\max}(\mathbf{C}')$, we can approximate the EDF of \mathbf{C} by

$$f_{\lambda(\mathbf{C})}(x) \approx \frac{n-1}{n} f_{\lambda(\mathbf{C}')} + \frac{1}{n} \delta(x - np^2), \quad (12)$$

where $\delta(\cdot)$ is the Dirac delta function. This model yields

$$\mu_{\lambda(\mathbf{C})}(k; m, n) \approx \frac{n-1}{n} \mu_{\lambda(\mathbf{C}')} + \frac{1}{n} (np^2)^k, \quad (13)$$

by the linearity of integration.

Evaluation of our trace-expectation factor requires finding $\mu_{\lambda(\mathbf{C})}(-1; m, n)$. Because $np^2 \gg q_+$, the second term in (13) is negligible in comparison to the first:

$$\mu_{\lambda(\mathbf{C})}(-1; m, n) \approx \frac{n-1}{n} \mu_{\lambda(\mathbf{C}')}(-1; m, n). \quad (14)$$

Combining (8), (11), and (14), we obtain the following analytical mse approximation in the constant q regime:

$$\text{mse-rmt} = \frac{p(n-1)}{m} \|\mathbf{x}\|_1 \mu_{\lambda(\mathbf{C}')}(-1; m, n). \quad (15)$$

Finally, using the Wishart-matrix result $\mu_{\lambda(\mathbf{C}')}(-1; m, n) = m/(\sigma^2(m-n))$ [15], [26], together with the Bernoulli variable's $\sigma^2 = p(1-p)$ variance, we arrive at our RMT approximation for the least-squares estimator's MSE:

$$\text{mse-rmt} = \frac{(n-1)\|\mathbf{x}\|_1}{(1-p)(m-n)}. \quad (16)$$

B. Numerical Experiments

To validate (16), we compared it to results obtained from numerical experiments. Also included was comparison with the baseline MSE approximation as follows. Using a law of large numbers, the approximation

$$\frac{\mathbf{A}^T \mathbf{A}}{m} \approx \lim_{m \rightarrow \infty} \frac{\sum_{k=1}^m \mathbf{a}_k \mathbf{a}_k^T}{m} = \mathbb{E}[\mathbf{a} \mathbf{a}^T] = p(1-p) \mathbf{I}_n + p^2 \mathbf{1}_{n \times n}, \quad (17)$$

where $\mathbf{1}_{n \times n}$ is the $n \times n$ matrix of ones, is asymptotically exact as the number of observations grows at fixed signal dimension, i.e., $m \rightarrow \infty$ with n fixed implying that $q \rightarrow 0$.

The inverse of the matrix in (17) is

$$\frac{1}{p(1-p)} \mathbf{I}_n - \frac{1}{(1-p)(np+1-p)} \mathbf{1}_{n \times n}, \quad (18)$$

which permits us to approximate (8) by

$$\begin{aligned} \text{mse-baseline} &= p \|\mathbf{x}\|_1 \text{Tr} \left(\frac{(\mathbb{E}[\mathbf{a} \mathbf{a}^T])^{-1}}{m} \right) \\ &= \frac{p \|\mathbf{x}\|_1}{m} \text{Tr} \left(\frac{1}{p(1-p)} \mathbf{I}_n - \frac{1}{(1-p)(np+1-p)} \mathbf{1}_{n \times n} \right) \\ &= \frac{p \|\mathbf{x}\|_1}{m} \left(\frac{n}{p(1-p)} - \frac{n}{(1-p)(np+1-p)} \right) \\ &= \frac{p \|\mathbf{x}\|_1}{m} \frac{n((n-2)p+1)}{p(1-p)(np+1-p)} \\ &= \frac{(n-2)p+1}{(n-1)p+1} \frac{n \|\mathbf{x}\|_1}{(1-p)m}. \end{aligned} \quad (19)$$

The first factor in (19) is approximately 1 for large n ; the second factor is approximately the same as (16) when $q \rightarrow 0$.

In Fig. 3, we compare the conventional closed-form approximation *mse-baseline* ($q \rightarrow 0$), the proposed closed-form approximation *mse-rmt* (constant q), and the *mse* computed from 500 Monte Carlo trials for each pair of the imaging parameters n and p . In these simulations, \mathbf{x} was drawn i.i.d. from the uniform distribution from 0 to 10. The Python simulation code we used is available from [28].

For moderate image dimension ($n = 100$), Fig. 3 shows that *mse-rmt* almost exactly matches *mse*, while the baseline asymptotic model *mse-baseline* is a lower bound on *mse* with a non-trivial gap for q values greater than 0.4. For the smaller image dimension ($n = 20$) with $q > 0.8$, we see that our RMT-based approximation, although better than *mse-baseline*, deviates from the true MSE. Such behavior is to be expected, because (6) presumed large problem size.

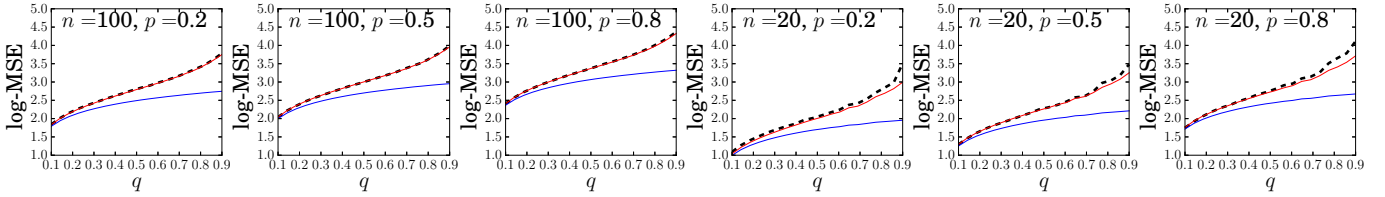


Fig. 3. Logarithmic (base-10) plots of mse in dashed black, mse-rmt in solid red, and mse-baseline in solid blue for different Bernoulli-distribution success probabilities $p \in \{0.2, 0.5, 0.8\}$ and signal dimensions $n \in \{20, 100\}$. Each log-MSE plot is versus $q = n/m$, the signal dimension divided by the number of observations. The total computational times for mse, mse-rmt, and mse-baseline respectively for all six plots were 181, 0.001, and 0.002 seconds on a MacBook pro with 2.2 GHz processor.

IV. ANALYSIS FOR IMAGING PARAMETER DESIGN

Because (16) is a good closed-form approximation to the MSE, it can be used to choose optimal acquisition parameters given imaging constraints. Suppose that the single-pixel camera is equipped with its own light source. Then (16) enables answering the following question: What is the minimum number of observations and the lowest incident flux that can guarantee that the MSE of the least-squares image lies below a prespecified value? Defining normalized root-mean-squared error (NRMSE) as $\text{nrmse-rmt} = \sqrt{\text{mse-rmt}}/\|\mathbf{x}\|_1$, our approximation shows that $\text{NRMSE} \propto 1/\sqrt{m\|\mathbf{x}\|_1}$, when $m \gg n$, i.e., it is inversely proportional to the square root of the number of observations, \sqrt{m} , times the square root of the total optical flux, $\sqrt{F} = \sqrt{\|\mathbf{x}\|_1}$. For example, suppose p and n are fixed. Then the set of minimal number of observations and optical flux, given a maximum tolerable normalized root-mean-squared error, is

$$\mathcal{S} = \{(m, F) : \sqrt{(n-1)/((1-p)(m-n)F)} = \text{nrmse-rmt}\}.$$

Figure 4 shows four least-squares-solution examples for the letter ‘R’ sample scene from [1]. These $n = 18 \times 18 = 324$ pixel images with $p = 0.5$ were generated by computer simulations. Three of them employed (m, F) values from \mathcal{S} with $\text{nrmse-rmt} = 0.01$, while the other employed an (m, F) pair far removed from \mathcal{S} . We see that the three solutions with (m, F) in \mathcal{S} have very similar image qualities, and they are clearly superior to the one whose (m, F) deviates from \mathcal{S} .

Note that (16) shows MSE decreasing monotonically with decreasing p . But a smaller p value implies a larger condition

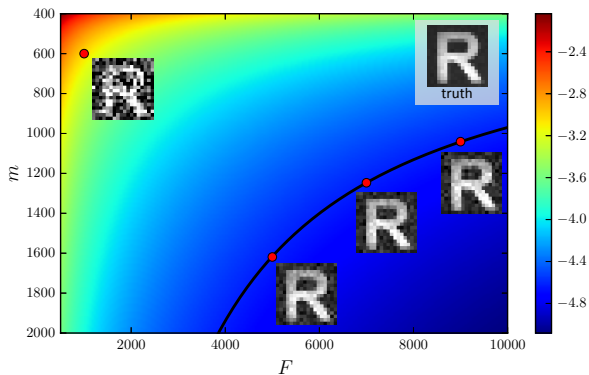


Fig. 4. Natural logarithms of nrmse-rmt (shown on the color bar) for four least-squares images of the letter ‘R’ sample scene (top right) that were generated from computer-simulated photon count data using Eq. (2) for different (m, F) pairs. Here, we used Bernoulli patterns for \mathbf{A} . Each image is shown immediately below the associated red dot marking its (m, F) pair. The black curve shows our \mathcal{S} for $\text{nrmse-rmt} = 0.01$ ($\log_e(\text{nrmse-rmt}) = -4.6$).

number for \mathbf{A} , because that matrix’s smallest singular value is quadratic in p , while (12) shows that its other $n - 1$ singular values are clustered around 1 as $p \rightarrow 0$. Since (16) assumes a well-defined $\hat{\mathbf{x}}_{\text{LS}}$, it fails to capture small p values leading to the occurrence of $\text{rank}(\mathbf{A}) < n$ with high probability. So we cannot conclude that reducing p will always reduce MSE.

V. CONCLUSIONS

In this letter, we used moment computation methods from random matrix theory to derive a remarkably simple closed-form approximation to the MSE of a low-flux least-squares single-pixel imaging system. Unlike conventional Monte Carlo methods (high accuracy, long computational time) and asymptotic approximation techniques (low accuracy, short computational time), our RMT-based MSE approximation achieves both high accuracy and high computational efficiency, as verified by numerical experiments. We also demonstrated how our framework can be used to understand limits of imaging parameter design given NRMSE constraints, and thus can inform the user when to consider improving fundamental system design by using higher efficiency detectors, for example.

It is of future interest to see how the proposed framework might be extended to other practical estimators for single-pixel imaging. For example, the Tikhonov estimator regularizes the least-squares estimate by perturbing the correlation matrix with a positive-definite matrix [29], and is useful for modeling smoothness image priors. A perturbation analysis on the correlation matrix might lead to an accurate MSE approximation for the Tikhonov estimator.

APPENDIX

The least-squares estimator’s error vector is

$$\mathbf{x} - \hat{\mathbf{x}}_{\text{LS}} = \mathbf{x} - (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T (\mathbf{A} \mathbf{x} + \boldsymbol{\eta}) = -(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\eta},$$

so the MSE can be written as

$$\begin{aligned} \text{mse} &= \mathbb{E} \left[\left((\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\eta} \right)^T (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\eta} \right] \\ &= \mathbb{E} \left[\boldsymbol{\eta}^T \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\eta} \right] \\ &\stackrel{(a)}{=} \mathbb{E} \left[\text{Tr} \left(\boldsymbol{\eta}^T \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T \boldsymbol{\eta} \right) \right] \\ &\stackrel{(b)}{=} \mathbb{E} \left[\text{Tr} \left(\boldsymbol{\eta} \boldsymbol{\eta}^T \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T \right) \right] \\ &\stackrel{(c)}{=} \text{Tr} \left(\mathbb{E} \left[\boldsymbol{\eta} \boldsymbol{\eta}^T \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-2} \mathbf{A}^T \right] \right), \end{aligned}$$

where: in (a) the trace of a scalar is the same scalar; in (b) the trace of a product is invariant to cyclic permutation of its factors; and in (c) the trace and expectation operations commute.

REFERENCES

- [1] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, 2008.
- [2] R. I. Stantchev, B. Sun, S. M. Hornett, P. A. Hobson, G. M. Gibson, M. J. Padgett, and E. Hendry, "Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector," *Science Advances*, vol. 2, no. 6, e1600190, 2016.
- [3] D. L. Snyder, *Random Point Processes*. Wiley, 1975.
- [4] A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, "First-photon imaging," *Science*, vol. 343, no. 6166, pp. 58–61, 2014.
- [5] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro, "Computational 3D and reflectivity imaging with high photon efficiency," in *Proc. IEEE Int. Conf. Image Process.*, 2014, pp. 46–50.
- [6] —, "Photon-efficient computational 3D and reflectivity imaging with single-photon detectors," *IEEE Trans. Comput. Imaging*, vol. 1, no. 2, pp. 112–125, Jun. 2015.
- [7] Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin, "Lidar waveform-based analysis of depth images constructed using sparse single-photon data," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 1935–1946, May 2016.
- [8] D. Shin, F. Xu, D. Venkatraman, R. Lussana, F. Villa, F. Zappa, V. K. Goyal, F. N. C. Wong, and J. H. Shapiro, "Photon-efficient imaging with a single-photon camera," *Nat. Commun.*, vol. 7, art. no. 12046, 24 Jun 2016.
- [9] W. Becker, A. Bergmann, M. A. Hink, K. König, K. Benndorf, and C. Biskup, "Fluorescence lifetime imaging by time-correlated single-photon counting," *Microscopy Res. Techn.*, vol. 63, no. 1, pp. 58–66, 2004.
- [10] M. Raginsky, R. M. Willett, Z. T. Harmany, and R. F. Marcia, "Compressed sensing performance bounds under Poisson noise," *Nat. Commun.*, vol. 6, 2015, doi: 10.1038/ncomms6913.
- [11] Z. T. Harmany, R. M. Willett, Z. T. Harmany, and R. F. Marcia, "This is SPIRAL-TAP: Sparse Poisson intensity reconstruction algorithms – theory and practice," *IEEE Trans. Image Proc.*, vol. 21, no. 3, pp. 1084–1096, 2012.
- [12] M. J. Sun, M. P. Edgar, D. B. Phillips, G. M. Gibson, and M. J. Padgett, "Improving the signal-to-noise ratio of single-pixel imaging using digital microscanning," *Opt. Expr.*, vol. 24, no. 10, pp. 10476–10485, 2016.
- [13] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*. Athena Scientific, 2002.
- [14] M. Pajovic and J. C. Preisig, "Performance analysis of the least squares based LTI channel identification algorithm using random matrix methods," in *Proc. 49th Ann. Allerton Conf. Commun., Control, Comput.*, 2011, pp. 516–523.
- [15] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Now Publishers, 2004.
- [16] D. C. Hoyle, "Accuracy of pseudo-inverse covariance learning—a random matrix theory analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 7, pp. 1470–1481, 2011.
- [17] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 1983–2010, Jun. 2005.
- [18] S. Rangan, A. Fletcher, and V. K. Goyal, "Asymptotic analysis of MAP estimation via the replica method and applications to compressed sensing," *IEEE Trans. Inform. Theory*, vol. 58, no. 3, pp. 1902–1923, Mar. 2012.
- [19] A. Gatti, M. Bache, D. Magatti, E. Brambilla, F. Ferri, and L. A. Lugiato, "Coherent imaging with pseudo-thermal incoherent light," *J. Mod. Optics*, vol. 53, no. 5–6, pp. 739–760, 2006.
- [20] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "A theory of multiplexed illumination," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, vol. 2, 2003, pp. 808–815.
- [21] —, "Multiplexing for optimal lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1339–1354, 2007.
- [22] N. Ratner, Y. Y. Schechner, and F. Goldberg, "Optimal multiplexed sensing: bounds, conditions and a graph theory link," *Opt. Expr.*, vol. 15, no. 25, pp. 17 072–17 092, 2007.
- [23] D. Shin, A. Kirmani, and V. K. Goyal, "Low-rate Poisson intensity estimation using multiplexed imaging," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process.*, Vancouver, Canada, May 2013, pp. 1364–1368.
- [24] J. H. Shapiro, "Computational ghost imaging," *Phys. Rev. A*, vol. 78, no. 6, art. 061802, 2008.
- [25] V. A. Marchenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Matematicheskii Sbornik*, vol. 114, no. 4, pp. 507–536, 1967.
- [26] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Dec. 1999.
- [27] S. S. Haykin, *Adaptive Filter Theory*, 5th ed. Pearson Higher Education, 2013.
- [28] "GitHub repository for single-pixel imaging analysis," <https://github.com/photon-efficient-imaging/rmt-for-ghost>.
- [29] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*. Winston, 1977.