



**HAL**  
open science

# Small-Noise Sensitivity Analysis of Locating Pulses in the Presence of Adversarial Perturbation

Meghna Kalra, Maxime Ferreira Da Costa, Kiryung Lee

► **To cite this version:**

Meghna Kalra, Maxime Ferreira Da Costa, Kiryung Lee. Small-Noise Sensitivity Analysis of Locating Pulses in the Presence of Adversarial Perturbation. 2024. hal-04489137

**HAL Id: hal-04489137**

**<https://hal.science/hal-04489137>**

Preprint submitted on 4 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Small-Noise Sensitivity Analysis of Locating Pulses in the Presence of Adversarial Perturbation

Meghna Kalra and Kiryung Lee

*Department of Electrical and Computer Engineering  
The Ohio State University  
Columbus, OH, USA*

Maxime Ferreira Da Costa

*CentraleSupélec  
Université Paris-Saclay  
Gif-sur-Yvette, France*

**Abstract**—A fundamental small-noise sensitivity analysis of spike localization in the presence of adversarial perturbations and arbitrary point spread function (PSF) is presented. The analysis leverages the local Lipschitz property of the inverse map from measurement noise to parameter estimate. In the small noise regime, the local Lipschitz constant converges to the spectral norm of the noiseless Jacobian of the inverse map. An interpretable upper bound in terms of the minimum separation of spikes, norms and flatness of the PSF and its derivative, and the distribution of spike amplitudes is provided. Numerical experiments highlighting the relevance of the theoretical bound as a proxy to the local Lipschitz constant and its dependence on the key attributes of the problem are presented.

**Keywords**— spike localization, local Lipschitz property, sensitivity analysis, adversarial noise.

## I. INTRODUCTION

The localization of spikes convolved with a known point spread function (PSF) arises in various applications, including imaging, array processing, radar, and communications. While this problem has a long history in signal processing [1], recent progress in super-resolution established theory and computational methods for the spike localization against or even below the Rayleigh resolution limit (e.g. [2]–[6]). Yet, most of the literature assumes the PSF is a Dirac distribution, or the noise is stochastic, typically modeled as Gaussian. However, there are many practical applications in which the target pulse signal is convolved with a non-Dirac PSF and is corrupted with a structured or an adversarial noise. This yields a gap between the classical theoretical guarantees on super-resolution and their applicability in real-world use cases. Given a fixed noise energy, structured noise can deteriorate the estimation performance more than random noise. For example, the spike injection framework considers spurious adversarial spikes consuming the entire noise energy. The error estimate on the parameters under this noise model has been shown to be related to the distance between the true and the faked signal components and can yield a significantly greater error estimate on the true signal parameter than random perturbation. This found interest in the context of physical layer security, where an agent transmitted over a wireless channel seeks to conceal her physical location from an eavesdropper. When the spike injection pattern is known to the receiver, the physical location can be preserved without degrading the decoding of legitimate parties [7]. In surface Electromyography (EMG) modality, the signal is given as the convolution of spikes representing neural activations and patient-dependent PSF, known as motor unit action potential (MUAP). In this application, structural noise arises in the form of cross-talk, constituting neural signals from neighboring muscles and

introducing additional peaks in the signal reading [8]. This peak can be misinterpreted as the activation of the muscle under observation and can cause the EMG acquisition algorithm to make a wrong inference.

**Contributions:** This paper establishes a fundamental analysis of the localization of spikes in the presence of non-Dirac PSF and adversarial noise. We are particularly interested in characterizing the phase transition between two regimes, characterized by the feasibility and unfeasibility of stable estimation. We will quantify the phase transition condition given in an interpretable form depending on the minimum separation of the spikes, PSF characteristics, and amplitude dynamic range. Our approach to the sensitivity analysis is based on the local Lipschitz property of the inverse map that maps an arbitrary noise instance to a least squares estimator. This local Lipschitz property describes how smoothly a function behaves in the vicinity of any point within its domain. Essentially, it implies that within any small neighborhood around a point, the change in the function’s output does not exceed a certain rate when compared to the change in the input. This framework enables the worst-case analysis in terms of the noise pattern. To the best of the authors’ knowledge, we have not found prior work on the fundamental analysis of spike localization in this worst-case scenario.

We focus on an asymptotic scenario of a small noise where the signal-to-noise ratio is sufficiently high. In this regime, the Lipschitz constant of the inverse map converges to the spectral norm of its noiseless Jacobian, which we use as a proxy to characterize the sensitivity of the model. We start by providing an analytical expression of the noiseless Jacobian of the inverse map. In order to better highlight its dependency on the key attribute of the problem, such as the minimum separation between the pulses, the shape of the PSF, and the dynamic range of the amplitudes, we derive in Theorem III.2 a novel simple and interpretable upper estimate of its spectral norm. To achieve this goal, we rely on the recent advances characterizing the spectrum of weighted Vandermonde matrices [9], [10] through the analytical properties of the extremal solution of generalized Beurling–Selberg type approximation problems [11]–[14].

The fundamental sensitivity analysis in this paper contrasts with the classical Cramér-Rao bound in the following perspectives. While the Cramér-Rao bound applies only to the class of unbiased estimators, most of the practical estimators for spike localization are biased. Furthermore, except for a few examples, including the additive white Gaussian model, the computation of the Cramér-Rao bound is not easy. On the contrary, our fundamental analysis provides a unique characterization in the worst-case noise model and applies regardless of the noise distribution even to adversarial noise cases.

**Mathematical Notations:** Vectors and matrices are written in small boldface  $\mathbf{a}$  and capital boldface  $\mathbf{A}$ , respectively.  $\mathbf{A}^*$  denotes the conjugate transpose of matrix  $\mathbf{A}$ . The notations  $\|\mathbf{A}\|$ ,  $\|\mathbf{A}\|_{\mathbb{F}}$ ,  $\|\mathbf{A}\|_{\infty \rightarrow 2}$  refer to the spectral norm, Frobenius norm, and the largest  $\ell_2$  norm of a column of a matrix  $\mathbf{A}$ . The maximum eigenvalue and singular value of  $\mathbf{A}$  are denoted  $\lambda_{\max}(\mathbf{A})$  or  $\sigma_{\max}(\mathbf{A})$ , respectively. The shorthand notation  $[n]$  to denote the set of integers  $\{1, \dots, n\}$ . The

MK and KL were supported by NSF CAREER Award CCF-1943201.

MF is with the Laboratory of Signals and Systems (L2S), CNRS. His work was supported by ANR: ANR-20-IDEES-0002.

number  $j$  is the basis of imaginary numbers,  $j^2 = -1$ . For  $T > 0$ , the  $K$ -dimensional torus is written  $\mathbb{T}^K \sim (\mathbb{R}/T\mathbb{Z})^K$ .

**Organization:** Section II lays the formulation of the problem statement and the signal model. Section III presents the main sensitivity analysis and presents the supporting lemmas and the main theorem. Section IV presents the sketch of the proof of the main results. Section V presents the validation for the theoretical analysis and finally we conclude with Section VI.

## II. PROBLEM STATEMENT

Consider the parameter estimation of pulses from multiple snapshots. The observed signal in the  $\ell$ th snapshot is written as

$$y_\ell(t) = \sum_{k=1}^K x_{k,\ell} g(t - \tau_k), \quad \ell \in [L] := \{1, \dots, L\}.$$

The signal consists of a superposition of  $K$  pulses of known point spread function (PSF)  $g$  centered at common locations  $\{\tau_k\}_{k=1}^K \subset \mathbb{T}$  with varying amplitudes across snapshots. The goal of the problem is to estimate the unknown locations in  $\boldsymbol{\tau} = [\tau_1, \dots, \tau_K]^T$  from the noisy Fourier transform coefficients of  $y_\ell$  sampled at frequencies in  $\{f_i = \frac{2i-N-1}{2T} : i \in [N]\}$  across  $L$  snapshots. Define  $\boldsymbol{\gamma} \mapsto \boldsymbol{\Phi}_\boldsymbol{\gamma} \in \mathbb{C}^{N \times K}$  by  $[\boldsymbol{\Phi}_\boldsymbol{\gamma}]_{i,k} = e^{-j2\pi\gamma_k f_i}$  where  $\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_K]^T \in \mathbb{T}^K$ . Let  $\mathbf{Y} \in \mathbb{C}^{N \times L}$  collect all measurements so that  $[\mathbf{Y}]_{i,\ell}$  denotes the Fourier transform  $\hat{y}_\ell(f) = \int_{-\infty}^{\infty} y_\ell(t) e^{-j2\pi f t} dt$  at frequency  $f_i$ , corrupted with additive noise, for  $i \in [N]$  and  $\ell \in [L]$ . Then  $\mathbf{Y}$  is compactly written as

$$\mathbf{Y} = \mathbf{G}\boldsymbol{\Phi}_\boldsymbol{\tau}\mathbf{X} + \mathbf{Z}$$

where  $\mathbf{G} = \text{diag}(\hat{g}(f_1), \dots, \hat{g}(f_N))$  is a diagonal matrix, where  $\hat{g}$  denotes the Fourier transform of  $g$ ,  $\mathbf{X} \in \mathbb{C}^{K \times L}$  satisfies  $[\mathbf{X}]_{k,\ell} = x_{k,\ell}$ , and  $\mathbf{Z} \in \mathbb{C}^{N \times L}$  denotes additive noise. We assume that  $\mathbf{G}$  has at least  $K$  nonzero diagonal entries. Then  $\mathbf{G}\boldsymbol{\Phi}_\boldsymbol{\tau}$  has full column rank.

We consider the least squares estimator given by

$$\underset{\boldsymbol{\gamma} \in \mathbb{T}^K, \boldsymbol{\Upsilon} \in \mathbb{C}^{K \times L}}{\text{minimize}} \quad \|\mathbf{Y} - \mathbf{G}\boldsymbol{\Phi}_\boldsymbol{\gamma}\boldsymbol{\Upsilon}\|_{\text{F}}^2$$

Given  $\boldsymbol{\gamma}$ , the optimal  $\boldsymbol{\Upsilon}$  is given by  $(\mathbf{G}\boldsymbol{\Phi}_\boldsymbol{\gamma})^\dagger \mathbf{Y}$ . Therefore, the same optimal estimator of  $\boldsymbol{\tau}$  is obtained by

$$\underset{\boldsymbol{\gamma} \in \mathbb{T}^K}{\text{minimize}} \quad \left\| \mathbf{P}_\boldsymbol{\gamma}^\perp \mathbf{Y} \right\|_{\text{F}}^2 \quad (1)$$

where  $\mathbf{P}_\boldsymbol{\gamma}^\perp := \mathbf{I}_N - \mathbf{G}\boldsymbol{\Phi}_\boldsymbol{\gamma}(\mathbf{G}\boldsymbol{\Phi}_\boldsymbol{\gamma})^\dagger$  denotes the projection onto the orthogonal complement of the column space of  $\mathbf{G}\boldsymbol{\Phi}_\boldsymbol{\gamma}$ . Let  $\hat{\boldsymbol{\tau}}$  denote a minimizer to (1). In the noise-free scenario ( $\mathbf{Z} = \mathbf{0}$ ) under a mild condition on  $N$  and  $\mathbf{X}$ , the estimate is uniquely determined as the ground-truth  $\boldsymbol{\tau}$ . The objective is to study the perturbation of the least-squares estimator as a function of worst-case additive noise  $\mathbf{Z}$ .

## III. LOCAL SENSITIVITY ANALYSIS

For the purpose of analysis, we assume up to a scaling of the pulse location and the PSF that  $T = 1$ . We present a fundamental sensitivity analysis that is based on the local Lipschitz property of the inverse map  $\boldsymbol{\psi}$ , which takes the noise perturbation  $\mathbf{Z}$  as input and outputs the least-squares estimate  $\boldsymbol{\tau}$ . To elucidate the dependence on additive noise  $\mathbf{Z}$ , we rewrite the optimization in (1) as

$$\underset{\boldsymbol{\gamma} \in \mathbb{T}^K}{\text{minimize}} \quad \ell(\boldsymbol{\gamma}, \mathbf{Z})$$

where

$$\ell(\boldsymbol{\gamma}, \mathbf{Z}) := \left\| \mathbf{P}_\boldsymbol{\gamma}^\perp (\mathbf{G}\boldsymbol{\Phi}_\boldsymbol{\tau}\mathbf{X} + \mathbf{Z}) \right\|_{\text{F}}^2. \quad (2)$$

Then there exists a nonlinear mapping  $\boldsymbol{\psi} : \mathbb{C}^{N \times L} \rightarrow \mathbb{T}^K$  given by

$$\boldsymbol{\psi}(\mathbf{Z}) \in \underset{\boldsymbol{\gamma}}{\text{argmin}} \ell(\boldsymbol{\gamma}, \mathbf{Z}).$$

The inverse map  $\boldsymbol{\psi}$  is *locally Lipschitz continuous* at  $\mathbf{0} \in \mathbb{C}^{N \times L}$  if there exists a neighborhood  $\mathcal{N}$  of  $\mathbf{0}$  and a constant  $C > 0$  such that

$$\frac{\|\boldsymbol{\psi}(\mathbf{Z}_1) - \boldsymbol{\psi}(\mathbf{Z}_2)\|_2}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|_{\text{F}}} \leq C, \quad \forall \mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{N} : \mathbf{Z}_1 \neq \mathbf{Z}_2. \quad (3)$$

The local Lipschitz constant  $L_{\boldsymbol{\psi}, \mathcal{N}}$  denotes the smallest  $C > 0$  satisfying (3) and is characterized via the spectral norm of the Jacobian matrix of  $\boldsymbol{\psi}$  as

$$L_{\boldsymbol{\psi}, \mathcal{N}} = \sup_{\mathbf{Z} \in \mathcal{N}} \|\nabla_{\iota(\mathbf{Z})} \boldsymbol{\psi}(\mathbf{Z})\|$$

with  $\iota : \mathbb{C}^{L \times N} \rightarrow \mathbb{R}^{2LN}$  defined by  $\iota(\mathbf{Z}) = \begin{bmatrix} \text{Re}(\text{vec}(\mathbf{Z})) \\ \text{Im}(\text{vec}(\mathbf{Z})) \end{bmatrix}$  where  $\text{vec}(\mathbf{Z}) \in \mathbb{C}^{LN}$  is obtained by stacking all columns of  $\mathbf{Z}$  vertically. Note that  $\iota$  is an isometric bijection.

For an arbitrary  $\mathbf{Z}$ , the Jacobian  $\nabla_{\iota(\mathbf{Z})} \boldsymbol{\psi}(\mathbf{Z})$  of the inverse map  $\boldsymbol{\psi}$  at  $\mathbf{Z}$  has entries given as multivariate polynomials in  $\iota(\mathbf{Z})$  where the order depends on  $N$ . Let  $\mathcal{N} \subset \mathbb{C}^{L \times N}$  be a Frobenius-norm ball of radius  $r > 0$ . Since  $\|\nabla_{\iota(\mathbf{Z})} \boldsymbol{\psi}(\mathbf{Z})\|$  is a continuous function in  $\iota(\mathbf{Z})$ , in the limit of  $r \rightarrow 0$ , the local Lipschitz constant  $L_{\boldsymbol{\psi}, \mathcal{N}}$  will converge to  $\|\nabla_{\iota(\mathbf{Z})} \boldsymbol{\psi}(\mathbf{0})\|$ . Therefore, one may deduce that  $\|\nabla_{\iota(\mathbf{Z})} \boldsymbol{\psi}(\mathbf{0})\|$  describes the asymptotic sensitivity of the estimation process in the small-noise regime.

The following provides a closed-form expression of  $\|\nabla_{\iota(\mathbf{Z})} \boldsymbol{\psi}(\mathbf{0})\|$ . It was derived by extending pre-existing results on the differentiation of pseudo-inverse and projection [15] to the case of complex-valued matrices. The statement can be verified with appropriate modification on the transpose operators.

**Lemma III.1.** *The Jacobian matrices in the right-hand side of (1) at  $\mathbf{z} = \mathbf{0}$  are written as*

$$\nabla_{\boldsymbol{\gamma}} \boldsymbol{q}(\boldsymbol{\tau}, \mathbf{0}) = 2\text{Re} \left( \overline{\mathbf{X}} \mathbf{X}^T \odot \boldsymbol{\Phi}_\boldsymbol{\tau}^* \boldsymbol{\Lambda}^* \mathbf{G}^* \mathbf{P}_\boldsymbol{\tau}^\perp \mathbf{G} \boldsymbol{\Lambda} \boldsymbol{\Phi}_\boldsymbol{\tau} \right) \quad (4a)$$

and

$$\nabla_{\iota(\mathbf{Z})} \boldsymbol{q}(\boldsymbol{\tau}, \mathbf{0}) = 2 \begin{bmatrix} \text{Re} \left( \mathbf{X}^T * \mathbf{P}_\boldsymbol{\tau}^\perp \mathbf{G} \boldsymbol{\Lambda} \boldsymbol{\Phi}_\boldsymbol{\tau} \right) \\ \text{Im} \left( \mathbf{X}^T * \mathbf{P}_\boldsymbol{\tau}^\perp \mathbf{G} \boldsymbol{\Lambda} \boldsymbol{\Phi}_\boldsymbol{\tau} \right) \end{bmatrix}^T \quad (4b)$$

where  $\odot$  and  $*$  respectively denote the Hadamard and Khatri-Rao products, and  $\boldsymbol{\Lambda} \in \mathbb{R}^{N \times N}$  is a diagonal matrix satisfying  $[\boldsymbol{\Lambda}]_{i,i} = -j2\pi f_i$  for  $i \in [N]$ .

As we illustrate in Section V, the zero-noise Jacobian norm  $\|\nabla_{\iota(\mathbf{Z})} \boldsymbol{\psi}(\mathbf{0})\|$  characterizes the model sensitivity depending on minimum separation and PSF characteristics. However, the complicated form in (4) does not explicitly explain how the sensitivity depends on these key attributes. Therefore, our main result derives an upper estimate of  $\|\nabla_{\iota(\mathbf{Z})} \boldsymbol{\psi}(\mathbf{0})\|$  in an interpretable form.

To state the main theorem, we introduce relevant notation. Let  $E_0$  and  $E_1$  be the bandlimited energy of the PSF and its derivative in the bandwidth  $J_N := [-\frac{N-1}{2}, \frac{N-1}{2}]$  so that

$$E_0 := \|\hat{g} \mathbb{1}_{J_N}\|_{L_2}^2, \quad E_1 := \|\hat{g}' \mathbb{1}_{J_N}\|_{L_2}^2, \quad (5)$$

where  $\mathbb{1}_{J_N}$  is the indicator function of the interval  $J_N$  and  $\hat{g}'$  denotes the Fourier transform of the first-order derivative of  $g$ . Given  $\mathcal{C}_0$  the space of continuous functions of the real variable, the total variation norm  $\|\cdot\|_{\text{TV}}$  of a measure  $q$  is defined as  $\|q\|_{\text{TV}} = \sup_{h \in \mathcal{C}_0} \int_{-\infty}^{\infty} h(f) dq(f)$ . In the sequel define  $\rho$  as the maximum of the normalized total variations of  $g$  and  $g'$  in the bandwidth  $J_N$ , i.e.

$$\rho := \max \left\{ \frac{\|\hat{g}\|^2 \mathbb{1}_{J_N}\|_{\text{TV}}}{E_0}, \frac{\|\hat{g}'\|^2 \mathbb{1}_{J_N}\|_{\text{TV}}}{E_1} \right\}. \quad (6)$$

Intuitively,  $\rho$  measures the ‘‘flatness’’ of the power spectral densities of  $g$  and  $g'$  within the interval  $J_N$ . It decreases as the PSF  $g$  and its derivative  $g'$  gets narrower in the time domain.

**Theorem III.2.** *Suppose that  $\min_{k \neq k'} \inf_{j \in Z} |\tau_k - \tau_{k'} + j| \geq \Delta$ . Let  $\kappa := \frac{\|\mathbf{X}\mathbf{X}^*\|_{\infty \rightarrow 2}}{\min(\text{diag}(\mathbf{X}\mathbf{X}^*))} \geq 1$  be the dynamic range of the spike amplitudes. If  $\Delta > \frac{2}{3}\rho\kappa$ , then*

$$\|\nabla_{\iota(\mathbf{z})}\psi(\mathbf{0})\| \leq \frac{\|\mathbf{X}\| \sqrt{1 + \frac{2}{3}\rho\Delta^{-1}}}{\sqrt{E_1} \min(\text{diag}(\mathbf{X}\mathbf{X}^*)) (1 - \frac{2}{3}\rho\kappa\Delta^{-1})}. \quad (7)$$

The parameter  $\Delta$  represents the *minimal separation* between any two pulse locations in the support; it has been known since the work of Rayleigh on diffraction to drive the harness of estimating  $\tau$  [16]. Additionally, when the spike amplitudes are drawn according to a statistical model, the dynamic range parameter  $\kappa$  is determined by the empirical covariance  $\mathbf{X}\mathbf{X}^*$  of the spike amplitudes. In the asymptotic number of snapshots,  $\mathbf{X}\mathbf{X}^*$  converges to the true covariance. When the amplitudes are uncorrelated,  $\kappa$  reduces to the dynamic range of variances of the spike amplitudes. In the other extreme case of a single snapshot,  $\kappa$  will be larger with nonzero off-diagonal entries in  $\mathbf{X}\mathbf{X}^*$ . In this perspective,  $\kappa$  also explains how multiple snapshots contribute to mitigating the sensitivity of the parameter estimation.

Note that the right-hand side of (7) is inversely proportional to the scaling of  $g$  and  $\mathbf{X}$ . Next, we present a corollary of Theorem III.2 and the following lemma, providing an appropriate normalization of the estimation error by the observation energy.

**Lemma III.3** ([17, Theorem 1]). *For any  $\Delta > 0$ , we have*

$$\sigma_{\max}(\mathbf{G}\Phi_{\tau}) \leq \sqrt{E_0 \left(1 + \frac{1}{2}\rho\Delta^{-1}\right)}.$$

To state our result, let us introduce  $\text{SNR} := \frac{\|\mathbf{G}\Phi_{\tau}\mathbf{X}\|_{\text{F}}^2}{\|\mathbf{z}\|_{\text{F}}^2}$ . Then in the limit of  $\text{SNR} \rightarrow \infty$ , we have

$$\|\psi(\mathbf{z}) - \tau\|_2 \leq \|\nabla_{\iota(\mathbf{z})}\psi(\mathbf{0})\| \cdot \|\mathbf{G}\Phi_{\tau}\mathbf{X}\|_{\text{F}} \cdot \text{SNR}^{-1/2}. \quad (8)$$

The next corollary presents a non-asymptotic upper bound on the noise propagation factor in (8).

**Corollary III.4.** *Suppose that the hypothesis of Theorem III.2 holds. Then*

$$\|\nabla_{\iota(\mathbf{z})}\psi(\mathbf{0})\| \cdot \|\mathbf{G}\Phi_{\tau}\mathbf{X}\|_{\text{F}} \leq \frac{\eta \sqrt{(1 + \frac{2}{3}\rho\Delta^{-1})(1 + \frac{1}{2}\rho\Delta^{-1})}}{\sqrt{\frac{E_1}{E_0}} (1 - \frac{2}{3}\rho\kappa\Delta^{-1})} \quad (9)$$

where  $\eta := \frac{\|\mathbf{X}\| \cdot \|\mathbf{X}\|_{\text{F}}}{\min(\text{diag}(\mathbf{X}\mathbf{X}^*))}$ .

Corollary III.4 describes how the adversarial noise propagates to the estimate by the inverse map  $\psi$ . The parameter  $\eta$  is another measure of the dynamic range of the spike amplitudes. Similar to  $\kappa$ , the asymptotic number of snapshots is determined by the true covariance matrix. When the amplitudes are uncorrelated,  $\eta$  is proportional to  $\sqrt{K}$ .

The result also suggests improved stability of the inverse map to small-noise for small values of  $\rho$ . Provided sufficient flatness of the power spectrum densities of  $g$  and  $g'$ , one would expect  $\rho$  to be inversely proportional to the bandwidth  $|J_N| = N$ , i.e.  $\rho = \mathcal{O}(N^{-1})$ . For example,  $\rho \simeq 8/N$  when  $g$  is a Dirac function.

## IV. PROOF OF THE MAIN RESULTS

### A. Proof of Theorem III.2

Let  $\mathbf{q}$  denote the partial gradient of the loss function  $\ell$  in (2) with respect to  $\gamma$ , i.e.  $\mathbf{q}(\gamma, \mathbf{z}) = \nabla_{\gamma}\ell(\gamma, \mathbf{z})$ . Then it has been shown [18]

that the *implicit function theorem* [19] (also see [18, Theorem 6]) can be utilized to express the Jacobian of the inverse map  $\psi$  as

$$\nabla_{\iota(\mathbf{z})}\psi(\mathbf{z}) = -(\nabla_{\gamma}\mathbf{q}(\psi(\mathbf{z}), \mathbf{z}))^{-1} (\nabla_{\iota(\mathbf{z})}\mathbf{q}(\psi(\mathbf{z}), \mathbf{z})). \quad (10)$$

From (10) and by the properties of the spectral norm, we have

$$\|\nabla_{\iota(\mathbf{z})}\psi(\mathbf{0})\| \leq \frac{\|\nabla_{\iota(\mathbf{z})}\mathbf{q}(\tau, \mathbf{0})\|}{\sigma_{\min}(\nabla_{\gamma}\mathbf{q}(\tau, \mathbf{0}))}. \quad (11)$$

To simplify the notation, we let in the sequel  $\mathbf{S}$  the positive definite Hermitian matrix  $\mathbf{S} = \Phi_{\tau}^* \Lambda^* \mathbf{G}^* \mathbf{P}_{\tau}^{\perp} \mathbf{G} \Lambda \Phi_{\tau}$ . We continue the proof by controlling the denominator and the numerator on the right-hand side of (11) separately in terms of the quantity  $\|\mathbf{S} - E_1 \mathbf{I}_K\|$  from their expression provided by Lemma III.1.

To provide a lower bound on the denominator on the right-hand side of (4a), we write

$$\begin{aligned} \frac{1}{2} \nabla_{\gamma}\mathbf{q}(\tau, \mathbf{0}) &= \text{Re}(\overline{\mathbf{X}}\mathbf{X}^{\top} \odot \mathbf{S}) \\ &= \text{Re}(\overline{\mathbf{X}}\mathbf{X}^{\top} \odot E_1 \mathbf{I}_K + \overline{\mathbf{X}}\mathbf{X}^{\top} \odot (\mathbf{S} - E_1 \mathbf{I}_K)) \\ &= E_1 \overline{\mathbf{X}}\mathbf{X}^{\top} \odot \mathbf{I}_K + \text{Re}(\overline{\mathbf{X}}\mathbf{X}^{\top} \odot (\mathbf{S} - E_1 \mathbf{I}_K)), \end{aligned}$$

where the last identity holds since  $\overline{\mathbf{X}}\mathbf{X}^{\top} \odot \mathbf{I}_K$  is a diagonal matrix with nonnegative entries. By [20, Theorem 1], we have

$$\begin{aligned} \left\| \text{Re}(\overline{\mathbf{X}}\mathbf{X}^{\top} \odot (\mathbf{S} - E_1 \mathbf{I}_K)) \right\| &\leq \left\| \overline{\mathbf{X}}\mathbf{X}^{\top} \odot (\mathbf{S} - E_1 \mathbf{I}_K) \right\| \\ &\leq \left\| \overline{\mathbf{X}}\mathbf{X}^{\top} \right\|_{\infty \rightarrow 2} \|\mathbf{S} - E_1 \mathbf{I}_K\|. \end{aligned}$$

By substituting the expression of  $\mathbf{A}$  and  $\mathbf{S}$ , we obtain

$$\begin{aligned} \frac{1}{2} \sigma_{\min}(\nabla_{\gamma}\mathbf{q}(\tau, \mathbf{0})) &\geq E_1 \min(\text{diag}(\mathbf{X}\mathbf{X}^*)) \\ &\quad - \|\mathbf{S} - E_1 \mathbf{I}_K\| \cdot \|\mathbf{X}\mathbf{X}^*\|_{\infty \rightarrow 2}. \quad (12) \end{aligned}$$

On the other hand, the numerator on the right-hand side of (11) can be upper-bounded with the inequality  $\|\mathbf{A} * \mathbf{B}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$  for any matrices  $\mathbf{A}$  and  $\mathbf{B}$  of the same number of columns. This yields

$$\begin{aligned} \frac{1}{2} \|\nabla_{\iota(\mathbf{z})}\mathbf{q}(\tau, \mathbf{0})\| &\leq \|\mathbf{X}\| \cdot \left\| \mathbf{P}_{\tau}^{\perp} \mathbf{G} \Lambda \Phi_{\tau} \right\| \\ &\leq \|\mathbf{X}\| \cdot \|\mathbf{S}\|_2^{1/2} \leq \|\mathbf{X}\| \cdot \sqrt{E_1 + \|\mathbf{S} - E_1 \mathbf{I}_K\|} \quad (13) \end{aligned}$$

Lemma IV.1 proposes a bound on the quantity  $\|\mathbf{S} - E_1 \mathbf{I}_K\|$  in terms of the separation between the pulses  $\Delta$  and the PSF-related parameters  $E_1$  and  $\rho$  defined in Equations (5) and (6). These bounds are derived via a generalization of the Beurling–Selberg approximation by Vaaler [11]. Its proof is deferred to Section IV-B for readability.

**Lemma IV.1.** *Assume that  $g(t) \in L_2(\mathbb{R})$  is derivable with  $g'(t) \in L_2(\mathbb{R})$  then for  $\Delta > \frac{2}{3}\rho$ , we have*

$$\|\mathbf{S} - E_1 \mathbf{I}_K\| \leq \frac{2}{3} E_1 \rho \Delta^{-1}. \quad (14)$$

One immediately concludes on the desired statement with Lemma IV.1, and by substituting Equations (12) and (13) into Equation (14). ■

### B. Proof of Lemma IV.1

Let  $\alpha = \sqrt{\frac{E_1}{E_0}}$  be a scaling factor, define  $\mathbf{U} = [\alpha \mathbf{G}\Phi_{\tau}, \Lambda \mathbf{G}\Phi_{\tau}]$ , and  $\mathbf{M} = \mathbf{U}^* \mathbf{U}$ . By a direct calculation, we have the block decomposition

$$\mathbf{M} = \begin{bmatrix} \alpha^2 \Phi_{\tau}^* \mathbf{G}^* \mathbf{G} \Phi_{\tau} & \alpha \Phi_{\tau}^* \mathbf{G}^* \Lambda \mathbf{G} \Phi_{\tau} \\ \alpha \Phi_{\tau}^* \mathbf{G}^* \Lambda^* \mathbf{G} \Phi_{\tau} & \Phi_{\tau}^* \mathbf{G}^* \Lambda^* \Lambda \mathbf{G} \Phi_{\tau} \end{bmatrix}.$$

By the linear independence of trigonometric polynomials and their derivatives, and since  $\mathbf{G}$  has at least  $2K$  non-zero diagonal entries, the

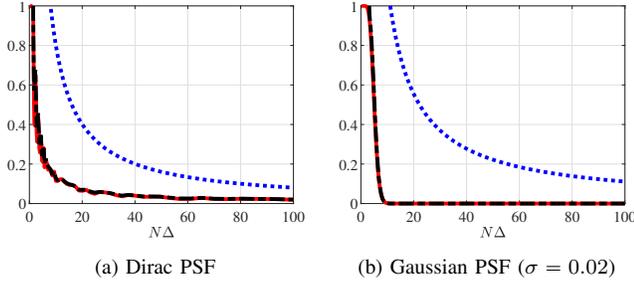


Figure 1. Empirical realization of the quantities  $\|\mathbf{S} - E_1 \mathbf{I}_K\|$  (in red/solid),  $\|\mathbf{M} - E_1 \mathbf{I}_K\|$  (in black/dash-dot), and of theoretical upper bound Equation (14) from Lemma IV.1 (in blue/dotted), for varying  $\Delta$ . Herein,  $N = 501$ .

matrices  $\mathbf{G}\Phi_\tau$ ,  $\Lambda\mathbf{G}\Phi_\tau$ , and  $\mathbf{U}$  are full column rank. This implies the matrix  $\mathbf{M}$ , and its two diagonal blocks are invertible. Hence, from the Schur block inversion formula, we have

$$\mathbf{M}^{-1} = \begin{bmatrix} * & * \\ * & \mathbf{S}^{-1} \end{bmatrix},$$

where we neglected the derivation of blocks marked with an asterisk. Since  $\mathbf{M}$ ,  $\mathbf{S}$  and their inverse are positive-definite, one may write

$$\lambda_{\max}(\mathbf{M}^{-1}) \geq \lambda_{\max}(\mathbf{S}^{-1}), \quad \lambda_{\min}(\mathbf{M}^{-1}) \leq \lambda_{\min}(\mathbf{S}^{-1}).$$

It comes with  $\lambda_{\min}(\mathbf{Q}^{-1}) = \lambda_{\max}(\mathbf{Q})^{-1}$ , and  $\lambda_{\max}(\mathbf{Q}^{-1}) = \lambda_{\min}(\mathbf{Q})^{-1}$  for any positive definite matrix  $\mathbf{Q}$  on the inequalities

$$\lambda_{\max}(\mathbf{M}) \geq \lambda_{\max}(\mathbf{S}), \quad \lambda_{\min}(\mathbf{M}) \leq \lambda_{\min}(\mathbf{S}). \quad (15)$$

With (15), we seek to bound the extremal eigenvalues of  $\mathbf{M}$ . Relevant bounds are provided in the following Lemma, recalled from [17] eigenvalues of  $\mathbf{M}$ , and rely on the Beurling–Selberg extremal approximation of functions with bounded variation.

**Lemma IV.2.** *For any  $\Delta > \frac{2}{3}\rho\kappa$ , one has the inequalities*

$$\lambda_{\min}(\mathbf{M}) \geq E_1(1 - \rho\Delta^{-1}), \quad \lambda_{\max}(\mathbf{M}) \geq E_1(1 + \rho\Delta^{-1}).$$

One concludes immediately with (15) and Lemma IV.2. ■

## V. NUMERICAL ILLUSTRATION

In this section, we numerically validate the effectiveness of our sensitivity analysis for the case of Dirac and Gaussian PSFs. The Gaussian PSF is given by  $g(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{t^2}{2\sigma^2}\right)$ , where the parameter  $\sigma > 0$  determines the effective width of the support. In this numerical illustration, we set the number of spikes to  $K = 3$  and the number of Fourier measurements to  $N = 501$ . The spike locations are generated randomly so that the minimum separation is no smaller than the parameter  $\Delta$ . The observation is maximized over 50 realizations in the Monte Carlo to simulate the worst-case scenario. The spike amplitudes are chosen at random and normalized to satisfy  $\frac{1}{L}\mathbf{X}\mathbf{X}^* = \mathbf{I}_K$ , which corresponds, provided the spike amplitudes are uncorrelated, to the asymptotic case of an infinity of snapshots goes to infinity.

Figure 1 compares empirical realization of the quantities  $\|\mathbf{S} - E_1 \mathbf{I}_K\|$  and  $\|\mathbf{M} - E_1 \mathbf{I}_K\|$  with their theoretical upper bound given by Equation (14) in Lemma IV.1. For both PSFs, the spectral distance between  $\mathbf{S}$  and the scaled identity  $E_1 \mathbf{I}_K$  improves with the time-bandwidth product  $N\Delta$ , suggesting better conditioning. This observation corroborates with the Lemma IV.1, which provides a valid upper bound for large enough values of  $N\Delta$ , and exhibits similar trends in the asymptotic. We also note that the empirical conditioning of the matrix  $\mathbf{S}$  and the established theoretical bound are higher for the Gaussian PSF than for the Dirac, the latter being a concentrated pulse with a lower spectral flatness factor  $\rho$ .

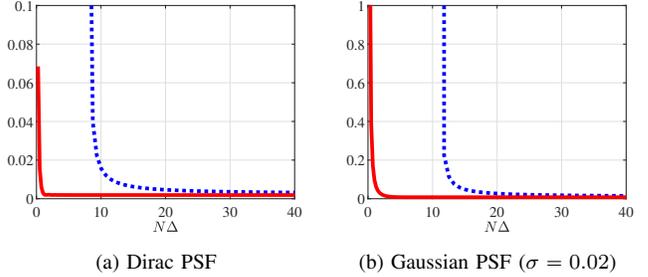


Figure 2. Empirical realization of the noise propagation factor  $\|\nabla_t(\mathbf{z})\psi(\mathbf{0})\| \cdot \|\mathbf{G}\Phi_\tau \mathbf{X}\|_F$  (in red/solid), and the upper bound (9) in Corollary III.4 (in blue/dotted) per varying  $\Delta$ .

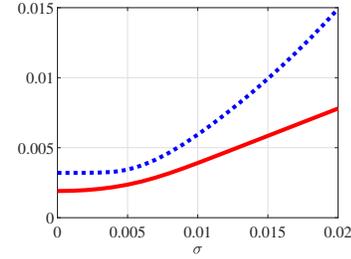


Figure 3. Illustration of the noise propagation factor (in red/solid) and its upper bound in Corollary III.4 (in blue/dotted) for the Gaussian PSF per varying  $\sigma$  ( $N = 501$ ,  $\Delta = 0.25$ ,  $N\Delta \approx 125$ ).

Next, we validate the statement of Corollary III.4 throughout numerical simulations. Figure 2 compares the noise propagation factor and its upper bound for both Dirac and Gaussian PSF. The empirical realization of the noise propagation factor (NPF) steeply increases when  $N\Delta$  is smaller than a threshold around 1, which corresponds to the Rayleigh resolution limit, indicating a highly sensitive regime. On the other hand, the NPF converges to a constant value as  $N\Delta$  increases, landing on the stable estimation regime. The dependence on the factor  $N\Delta$  for the stability of pulse-localization has been widely studied for stochastic noise. This experiment illustrates a similar phenomenon for adversarial noise. Meanwhile, the bound by Corollary III.4 provides a valid upper bound for sufficiently large  $N\Delta$ . It is also larger than the target quantity by a constant factor. Besides this conservativeness, the upper bound reflects the overall trend of NPG and is useful since it provides a simple interpretable relation to minimum separation, PSF characteristics, and spike amplitudes. Figure 3 pictures the NPF and its theoretical upper bound as a function of the width  $\sigma$  of a Gaussian PSF. It confirms the intuition that wider convolution kernels are more sensitive to adversarial noise.

## VI. CONCLUSION

This paper presents the fundamental sensitivity analysis of the spike localization problem in the presence of adversarial noise and arbitrary PSF. The analysis is based on the local Lipschitz property of the inverse map that controls the effect of the measurement noise on parameter estimation error and quantifies the high and low noise sensitivity regions. We focused on the high-SNR scenario where the local Lipschitz converges to the spectral norm of the noiseless Jacobian of the inverse map. This setting serves as a baseline measure of stability. Our main result derives an interpretable upper bound on this quantity in terms of the minimum separation, the shape of the PSF, and the spike amplitude distribution. The numerical section bolsters the efficacy of the derived bounds.

## REFERENCES

- [1] S. M. Kay and S. L. Marple, "Spectrum analysis—a modern perspective," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1380–1419, 1981.
- [2] G. Schiebinger, E. Robeva, and B. Recht, "Superresolution without separation," *Information and Inference: A Journal of the IMA*, vol. 7, no. 1, pp. 1–30, 2018.
- [3] B. Diederichs, "Well-posedness of sparse frequency estimation," *arXiv preprint arXiv:1905.08005*, 2019.
- [4] Y. Chi and M. Ferreira Da Costa, "Harnessing sparsity over the continuum: Atomic norm minimization for superresolution," *IEEE Signal Processing Magazine*, vol. 37, no. 2, pp. 39–57, 2020.
- [5] Q. Li and G. Tang, "Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision," *Applied and Computational Harmonic Analysis*, vol. 48, no. 3, pp. 891–948, 2020.
- [6] W. Li and W. Liao, "Stable super-resolution limit and smallest singular value of restricted Fourier matrices," *Applied and Computational Harmonic Analysis*, vol. 51, pp. 118–156, 2021.
- [7] J. Li and U. Mitra, "Channel state information-free location-privacy enhancement: Fake path injection," *arXiv preprint arXiv:2307.05442*, 2023.
- [8] C. J. De Luca, M. Kuznetsov, L. D. Gilmore, and S. H. Roy, "Inter-electrode spacing of surface EMG sensors: Reduction of crosstalk contamination during voluntary contractions," *Journal of biomechanics*, vol. 45, no. 3, pp. 555–561, 2012.
- [9] A. Moitra, "Super-resolution, extremal functions and the condition number of vandermonde matrices," in *ACM STOC*, 2015, pp. 821–830.
- [10] C. Aubel and H. Bölcskei, "Vandermonde matrices with nodes in the unit disk and the large sieve," *Applied and Computational Harmonic Analysis*, vol. 47, no. 1, pp. 53–86, 2019.
- [11] J. D. Vaaler, "Some extremal functions in Fourier analysis," *Bulletin of the American Mathematical Society*, vol. 12, no. 2, pp. 183–216, 2 1985.
- [12] E. Carneiro, "A survey on Beurling-Selberg majorants and some consequences of the Riemann hypothesis," *Matemática Contemporânea*, vol. 40, no. 8, 2011.
- [13] E. Carneiro and F. Littmann, "Extremal functions in de Branges and Euclidean spaces," *Advances in Mathematics*, vol. 260, pp. 281–349, Aug. 2014.
- [14] J. D. Vaaler, "On the number of lattice points in a ball." arXiv: 2303.14905 [math]. (May 24, 2023), [Online]. Available: <http://arxiv.org/abs/2303.14905>, preprint.
- [15] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal on numerical analysis*, vol. 10, no. 2, pp. 413–432, 1973.
- [16] J. Lindberg, "Mathematical concepts of optical superresolution," *Journal of Optics - IOP Publishing*, vol. 14, no. 8, p. 83 001, 2012.
- [17] M. Ferreira Da Costa, "The condition number of weighted non-harmonic Fourier matrices with applications to super-resolution." hal: hal-04261330. (Oct. 2023), preprint.
- [18] S. Basu and Y. Bresler, "The stability of nonlinear least squares problems and the Cramér-Rao bound," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3426–3436, 2000.
- [19] W. Rudin, *Principles of Mathematical Analysis*. McGraw-Hill, 1976.
- [20] X. Zhan, "Inequalities for the singular values of Hadamard products," *SIAM Journal on Matrix Analysis and Applications*, vol. 18, no. 4, pp. 1093–1095, 1997.