Convex Quantization Preserves Logconcavity

Pol del Aguila Pla, *Member, IEEE*, Aleix Boquet-Pujadas, and Joakim Jaldén, *Senior Member, IEEE*

***Abstract*—A logconcave likelihood is as important to proper statistical inference as a convex cost function is important to variational optimization. Quantization is often disregarded when writing likelihood models, ignoring the limitations of the physical detectors used to collect the data. These two facts call for the question: would including quantization in likelihood models preclude logconcavity? are the true data likelihoods logconcave? We provide a general proof that the same simple assumption that leads to logconcave continuous-data likelihoods also leads to logconcave quantized-data likelihoods, provided that convex quantization regions are used.**

***Index Terms*—Bayesian statistics, likelihood, privacy-aware data analysis, 1-bit compressed sensing, inverse problems.**

I. Introduction

INFERENCE from signals in the digital domain is of central importance in digital signal processing. However, the discrete nature of measurement devices is often disregarded or misrepresented in building data likelihood models. When quantization is deemed fine enough, the established procedure is to ignore it or to model it as additive noise [1], [2]. Conversely, the few works investigating coarse quantization do so under simplifying assumptions [3], [4], [5], suggesting that optimal estimation might be intractable. In this paper, we show that the exact likelihood of quantized data remains logconcave under widely applicable assumptions.

This result is of interest to several signal processing domains. For example, coarse quantization has been increasingly popular due to the advent of 1-bit compressed sensing techniques [3], [4], [5], [6], [7], [8], [9], which can be seen as its limiting case. These promise to incorporate low-cost high-speed analog-to-digital converters (ADC) into wireless-communications pipelines. Quantization is also influential in the privacy-enhancement literature [10], [11]. Besides simple noise-addition mechanisms, coarse quantization or "aggregation" is one of the most straightforward techniques to induce differential privacy and k-anonymity on a database. However,

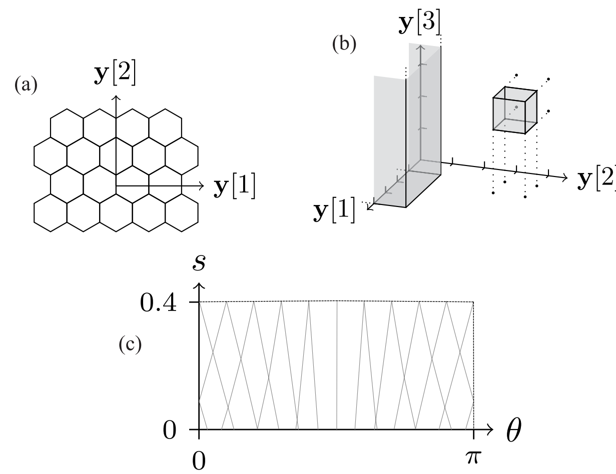

Fig. 1. a) Convex regions laid down according to a honeycomb pattern, as done in hexagonal electrode detectors. b) Two quantization regions of a convex quantizer formed by independent ADCs. c) Example quantization regions of a PET scanner in the sinogram domain.

privacy protection and data utility are conflicting objectives. To better understand this tradeoff in aggregation-based techniques, it is fundamental to characterize the properties of the likelihood after quantization.

Variational inverse problems with convex data terms, which can be interpreted as maximum-a-posteriori (MAP) estimators for logconcave likelihood models, also typically ignore data quantization in their cost functions. However, examples where quantization plays a crucial role abound in applications (see Fig. 1): think of honeycomb electrode detectors [12] in high-energy physics, the sinogram-domain quantization induced by the geometry of positron emission tomography (PET) scanners, the pixelized detectors in computed tomography [13], or CCD and CMOS photodetectors in biological microscopy [14]. Furthermore, we expect that Bayesian image reconstruction algorithms will benefit even further from treating quantization explicitly because they use the entire posterior distribution to make inferences. This is in contrast to using only the mean or the mode, which are much more easily preserved between continuous and quantized data.

Admittedly, incorporating quantization into models without some guarantee of likelihood logconcavity precludes many theoretical guarantees. Examples include tractable maximum likelihood (ML) and MAP estimates, connected and convex credible and confidence sets, as well as large-sample normal approximations that lead to usable hypothesis testing [15]. These are automatic for most common noise models (e.g., Gaussian or exponential) in the continuous-data setting. To recover these properties, most existing works studying quantization resort to

Manuscript received 18 August 2022; revised 3 December 2022; accepted 17 December 2022. Date of publication 29 December 2022; date of current version 13 January 2023. This work was partially supported by the SRA ICT TNG project Privacy-preserved Internet Traffic Analytics (PITA). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zubair Khalid. *(Corresponding author: Pol del Aguila Pla.)*

Pol del Aguila Pla is with the CIBM Center for Biomedical Imaging, 1015 Lausanne, Switzerland, and also with the Biomedical Imaging Group, École polytechnique fédérale de Lausanne, 1015 Lausanne, Switzerland (e-mail: pol.delaguilapla@epfl.ch).

Aleix Boquet-Pujadas is with Biomedical Imaging Group, École polytechnique fédérale de Lausanne, 1015 Lausanne, Switzerland (e-mail: aleix.boquetipujadas@epfl.ch).

Joakim Jaldén is with the Division of Information Science and Engineering at the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden (e-mail: jalden@kth.se).

Digital Object Identifier 10.1109/LSP.2022.3233001

simplifying assumptions such as Gaussian data and 1-bit quantization [3], [4], [5], or optimistically assume logconcavity [6], [9], [11]. In contrast, we prove that the logconcavity of the quantized-data likelihood follows from the same assumption made on its continuous-data counterpart: logconcavity of the noise distribution. The result holds for convex quantizers, which include not only all the examples mentioned so far, but also an overwhelming majority of applications. This means that the corresponding likelihood-based methods automatically enjoy convergence guarantees. Therefore, more accurate models can be readily adopted for a plethora of detector models. We expect that the convenience of this result will encourage researchers to more often include quantized models and thus better capture the nature of measurement devices. Similarly, another aim of this letter is to publicize and exploit a series of techniques that are part of the folklore of the statistics literature but are yet to be made known to the signal processing community [16]. For example, while not entirely new to our community [17], [18], [19], Prékopa's results [20] on logconcavity are key to this work and have not previously been harnessed in their full generality.

To the best of our knowledge, our analysis is the most general treatment of likelihood logconcavity for quantized data yet. Our study also accounts for the scale parameter, which has only been considered before in a specific application in [4]. Furthermore, we appear to be the first to do the analysis for a generic vector quantizer Q, which is more general than simple combinations of independent ADCs.

In Section II, we introduce the topic of likelihood logconcavity by studying it for continuous data models. In Section III, we present our main result: quantizers with convex quantization regions (convex quantizers) yield logconcave likelihoods when the underlying noise model has a logconcave pdf. In Section IV and the Appendix, we provide the proof for this statement, as well as supporting mathematical results.

II. Logconcavity for Continuous Data

Our central claim is that for both continuous and quantized data, likelihood logconcavity follows from the logconcavity of the noise pdf. This result is known for continuous data. We present it in Proposition 2.

We start by introducing the context and notation. Consider a data vector $\mathbf{y} \in \mathbb{R}^n$ modeled as

$$\mathbf{y} = \mathbf{\Psi}^{-1}\left(\mathbf{S}\mathbf{x} + \mathbf{w}\right), \tag{1}$$

where $\mathbf{\Psi} \in \mathcal{M}_n^+(\mathbb{R})$, $\mathbf{S} \in \mathcal{M}_{n,m}(\mathbb{R})$, and $\mathbf{x} \in \mathbb{R}^m$, with $\mathcal{M}_n^+(\mathbb{R})$ representing the set of $n \times n$ real positive-definite matrices and $\mathcal{M}_{n,m}(\mathbb{R})$ the space of $n \times m$ real matrices. Additionally, $\mathbf{w} \in \mathbb{R}^n$ is a random noise vector $\mathbf{w} \sim \mathrm{f}_{\mathbf{w}}(\mathbf{w})$ drawn from the logconcave pdf $\mathrm{f}_{\mathbf{w}}(\cdot)$, *i.e.*, for any $\alpha \in [0,1]$ and any two $\mathbf{w}_0, \mathbf{w}_1 \in \mathbb{R}^n$,

$$\mathrm{f}_{\mathbf{w}}(\mathbf{w}_\alpha) \geq \mathrm{f}_{\mathbf{w}}(\mathbf{w}_1)^\alpha\, \mathrm{f}_{\mathbf{w}}(\mathbf{w}_0)^{(1-\alpha)}. \tag{2}$$

Here, $\mathbf{w}_\alpha := \alpha\mathbf{w}_1 + (1-\alpha)\mathbf{w}_0$ is a convex combination between $\mathbf{w}_0$ and $\mathbf{w}_1$. Throughout the paper, we will resort to this notation for convex combinations for simplicity of exposition. The case $\mathbf{\Psi} = \mathbf{I}$ in (1) corresponds to a usual linear model formulation, where $\mathbf{S}$ is the observation matrix. In statistics, $\mathbf{x}$ and $\mathbf{\Psi}$ in (1) are known as the location and scale parameters of the distribution family defined by (1). For example, if the noise comes from a standard multivariate normal distribution $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, we have that $\mathbf{y} \sim \mathcal{N}(\mathbf{\Psi}^{-1}\mathbf{S}\mathbf{x}, \mathbf{\Psi}^{-2})$.

Remark 1 (Parametrization of Location-Scale Statistical Models): Although the location-scale parametrization in (1) is less familiar than, say, the mean and covariance matrices in multivariate Gaussian models, this choice is essentially related to the study of logconcavity. In fact, to show that joint estimators of $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ in a $\mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$ model are tractable, one needs to first reparametrize the problem in terms of $\mathbf{x}$ and $\mathbf{\Psi}$.

For each vector $\mathbf{y} \in \mathbb{R}^n$, the likelihood is a function of $\mathbf{x}$ and $\mathbf{\Psi}$ such that $\mathcal{L}(\cdot,\cdot;\mathbf{y}) : \mathbb{R}^m \times \mathcal{M}_n^+(\mathbb{R}) \to \mathbb{R}_+$ and

$$\mathcal{L}(\mathbf{x}, \mathbf{\Psi}; \mathbf{y}) := \mathrm{f}_{\mathbf{y};\mathbf{x},\mathbf{\Psi}}(\mathbf{y}) = \mathrm{f}_{\mathbf{w}}(\mathbf{\Psi}\mathbf{y} - \mathbf{S}\mathbf{x})\,. \tag{3}$$

Here, we used (1) to write $\mathcal{L}(\mathbf{x}, \mathbf{\Psi}; \mathbf{y})$ in terms of $\mathrm{f}_{\mathbf{w}}(\cdot)$, and $\mathrm{f}_{\mathbf{y};\mathbf{x},\mathbf{\Psi}}(\mathbf{y})$ denotes the pdf of $\mathbf{y}$ for some given values of $\mathbf{x}$ and $\mathbf{\Psi}$. The following result sets the stage for our study.

Proposition 2 (Logconcave Noise Generates Logconcave Likelihoods): Consider a sample $\mathbf{y} \in \mathbb{R}^n$ drawn from model (1) and assume that the pdf of the noise, $\mathrm{f}_{\mathbf{w}}(\cdot)$, is logconcave. Then, $\mathcal{L}(\mathbf{x}, \mathbf{\Psi}; \mathbf{y})$ is jointly logconcave in $\mathbf{x}$ and $\mathbf{\Psi}$.

Proof: Using (3), we obtain

$$\begin{aligned}\mathcal{L}(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha; \mathbf{y}) &= \mathrm{f}_{\mathbf{w}}(\mathbf{\Psi}_\alpha\mathbf{y} - \mathbf{S}\mathbf{x}_\alpha)\\ &\geq \mathrm{f}_{\mathbf{w}}([\mathbf{\Psi}_1\mathbf{y} - \mathbf{S}\mathbf{x}_1])^\alpha\, \mathrm{f}_{\mathbf{w}}([\mathbf{\Psi}_0\mathbf{y} - \mathbf{S}\mathbf{x}_0])^{1-\alpha},\end{aligned}$$

which is the desired result. Here, we used (2) with $\mathbf{w}_i = \mathbf{\Psi}_i\mathbf{y} - \mathbf{S}\mathbf{x}_i$ for $i \in \{0, 1\}$. □

III. Logconcavity for Quantized Data

It turns out that a statement similar to Proposition 2 can be made for quantized observations z modeled as

$$z = \mathrm{Q}\left(\mathbf{y}\right) = \mathrm{Q}\left(\mathbf{\Psi}^{-1}\left(\mathbf{S}\mathbf{x} + \mathbf{w}\right)\right). \tag{4}$$

Here, we take the most general view of quantization: we define a quantizer as a mapping $\mathrm{Q} : \mathbb{R}^n \to \mathcal{Z}$, where $\mathcal{Z}$ is a countable set. Such a quantizer does not generally treat each dimension of $\mathbf{y}$ independently. The results we present hereafter apply to the subclass of quantizers that have convex quantization regions. We call these *convex quantizers* Q, and they fulfill that $\mathrm{Q}^{-1}(z)$ is a convex set $\forall z \in \mathcal{Z}$ (see Fig. 1 for examples). Among others introduced beforehand, these include quantizers composed of independent (monotonic) ADCs for each dimension. Indeed, in this case, for any $z \in \mathcal{Z}$ and all $j \in \{1, 2, \ldots, n\}$, there are $a_j(z), b_j(z) \in \overline{\mathbb{R}}$ such that $\mathrm{Q}^{-1}(z) = \prod_{j=1}^n [a_j(z), b_j(z)]$, which is trivially convex. Here, $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. Our main result below parallels Proposition 2 for quantized data obtained from convex quantizers.

Theorem 3 (Logconcave Noise and Convex Quantizers Generate Logconcave Likelihoods): Consider a sample $z \in \mathcal{Z}$ drawn from model (4). Assume that Q is a convex quantizer and that the pdf $\mathrm{f}_{\mathbf{w}}(\cdot)$ of the noise is logconcave. Then,

1) for a given scale parameter $\mathbf{\Psi}_0 \in \mathcal{M}_n^+(\mathbb{R})$, the likelihood $\mathcal{L}(\mathbf{x}, \mathbf{\Psi}_0; z)$ is logconcave with respect to $\mathbf{x}$,
2) for scale parameters of the form $\mathbf{\Psi} = \psi\mathbf{I}$ with $\psi > 0$, the likelihood $\mathcal{L}(\mathbf{x}, \psi\mathbf{I}; z)$ is jointly logconcave with respect to $\mathbf{x}$ and ψ,
3) for diagonal positive-definite scale parameters $\mathbf{\Psi} = \mathbf{\Lambda}$, the likelihood $\mathcal{L}(\mathbf{x}, \mathbf{\Lambda}; z)$ is jointly logconcave with respect to $\mathbf{x}$ and $\mathbf{\Lambda}$ if Q is composed of independent ADCs for each dimension.

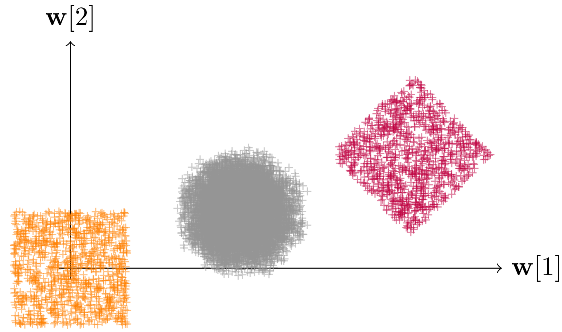

Fig. 2. Example of the Minkowski sum of two sets scaled by $\alpha = 1/2$ (in the center). Each set is represented here by random elements within it.

The rest of this letter is dedicated to proving Theorem 3. The proof can be found at the end of the letter.

IV. Mathematical Study of Logconcavity Under Quantization

We start by constructing the likelihood of the location and scale parameters for a given $z \in \mathcal{Z}$. By definition, observing z under model (4) implies that $\mathbf{y} \in \mathrm{Q}^{-1}(z)$. In turn, this implies that the random noise $\mathbf{w}$ is within a specific region. For specific values of $\mathbf{x}$ and $\mathbf{\Psi}$, and for each $z \in \mathcal{Z}$, we define

$$\mathcal{W}_z(\mathbf{x}, \mathbf{\Psi}) := \left\{\mathbf{w} \in \mathbb{R}^n : \mathbf{y} \in \mathrm{Q}^{-1}(z)\right\} \tag{5}$$

as a shorthand for this region. We then write the likelihood as

$$\mathcal{L}(\mathbf{x}, \mathbf{\Psi}; z) = \mathrm{P}_{\mathbf{w}}\left[\mathcal{W}_z(\mathbf{x}, \mathbf{\Psi})\right], \tag{6}$$

where $\mathrm{P}_{\mathbf{w}}$ is the probability law in $\mathbb{R}^n$ given by the pdf $\mathrm{f}_{\mathbf{w}}(\cdot)$.

Our strategy to prove Theorem 3 combines this compact expression of the quantized-data likelihood with a key result by Prékopa, which we include here for completeness.

Theorem 4 (Prékopa's Theorem [20, p. 2, Theorem 2]): Let $\mathbf{w}$ be a continuous random vector in $\mathbb{R}^n$ with logconcave pdf $\mathrm{f}_{\mathbf{w}}(\cdot)$ in the sense of (2). Let $\mathrm{P}_{\mathbf{w}} : 2^{\mathbb{R}^n} \to [0, 1]$ be the probability measure induced by $\mathbf{w}$ on $\mathbb{R}^n$, where $2^{\mathbb{R}^n}$ denotes the set of all possible sets in $\mathbb{R}^n$. Then, for any two convex sets $\mathcal{A}_0, \mathcal{A}_1 \subseteq \mathbb{R}^n$ we have that

$$\mathrm{P}_{\mathbf{w}}\left[\mathcal{A}_\alpha\right] \geq \mathrm{P}_{\mathbf{w}}\left[\mathcal{A}_1\right]^\alpha \mathrm{P}_{\mathbf{w}}\left[\mathcal{A}_0\right]^{(1-\alpha)}, \tag{7}$$

where $\mathcal{A}_\alpha$ is the Minkowski sum $\alpha\mathcal{A}_1 + (1-\alpha)\mathcal{A}_0$.

Here, the weighted Minkowski sum $\mathcal{A}_\alpha$ (illustrated in Fig. 2 for $\alpha = 1/2$) is the set of all possible combinations $\mathbf{w}_\alpha = \alpha\mathbf{w}_1 + (1-\alpha)\mathbf{w}_0$ in which $\mathbf{w}_1 \in \mathcal{A}_1, \mathbf{w}_0 \in \mathcal{A}_0$, for $\alpha \in [0, 1]$. The Minkowski sum preserves convexity: if $\mathcal{A}_1$ and $\mathcal{A}_2$ are convex, then $\mathcal{A}_\alpha$ is also convex.

To prove Theorem 3, we identify the sets $\mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha)$ in (5) with the sets $\mathcal{A}_\alpha$ in Theorem 4. Then, (7) becomes the desired logconcavity statement. The technical conditions of Theorem 3, therefore, only ensure that the convex combination of location and scale parameters leads to the same set $\mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha)$ as the Minkowski sum of the corresponding scaled sets $\mathcal{W}_z(\mathbf{x}_i, \mathbf{\Psi}_i)$ for $i \in \{0, 1\}$. We start by verifying convexity in the extreme cases $\alpha \in \{0, 1\}$.

Lemma 5 (Convex Quantizers Lead to Convex Noise Regions): For any $z \in \mathcal{Z}, \mathbf{x} \in \mathbb{R}^m$, and $\mathbf{\Psi} \in \mathcal{M}_n^+(\mathbb{R})$, $\mathcal{W}_z(\mathbf{x}, \mathbf{\Psi})$ is convex if and only if $\mathrm{Q}^{-1}(z)$ is convex.

Proof: Let $\mathbf{w}_0, \mathbf{w}_1 \in \mathcal{W}_z(\mathbf{x}, \mathbf{\Psi})$. Then, $\mathbf{w}_i = \mathbf{\Psi}\mathbf{y}_i - \mathbf{S}\mathbf{x}$ for $\mathbf{y}_i \in \mathrm{Q}^{-1}(z)$ with $i \in \{0, 1\}$. Because $\mathrm{Q}^{-1}(z)$ is convex, $\mathbf{y}_\alpha \in \mathrm{Q}^{-1}(z)$, and therefore, $\mathbf{w}_\alpha = \mathbf{\Psi}\mathbf{y}_\alpha - \mathbf{S}\mathbf{x}_\alpha \in \mathcal{W}_z(\mathbf{x}, \mathbf{\Psi})$. In conclusion, if $\mathrm{Q}^{-1}(z)$ is convex, $\mathcal{W}_z(\mathbf{x}, \mathbf{\Psi})$ is convex. □

For the converse, simply consider that $\mathcal{W}_z(\mathbf{0}, \mathbf{I}) = \mathrm{Q}^{-1}(z)$.

This allows us to set

$$\mathcal{A}_i := \mathcal{W}_z(\mathbf{x}_i, \mathbf{\Psi}_i) \text{ for } i \in \{0, 1\}, \tag{8}$$

while fulfilling the conditions of Theorem 4. For the intermediate values $\alpha \in (0, 1)$, we identify conditions under which $\mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha)$ is the Minkowski sum of $\alpha\mathcal{A}_1$ and $(1-\alpha)\mathcal{A}_0$. We start by showing that one of the inclusions is always true.

Lemma 6: Consider $\mathcal{A}_0$ and $\mathcal{A}_1$ as defined in (8). Then $\mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha) \subseteq \mathcal{A}_\alpha$.

Proof: Let $\mathbf{w} \in \mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha)$. Then, there is $\mathbf{y} \in \mathrm{Q}^{-1}(z)$ such that $\mathbf{w} = \mathbf{\Psi}_\alpha\mathbf{y} - \mathbf{S}\mathbf{x}_\alpha = \alpha\mathbf{w}_1 + (1-\alpha)\mathbf{w}_0$ with $\mathbf{w}_i = \mathbf{\Psi}_i\mathbf{y} - \mathbf{S}\mathbf{x}_i$ for $i \in \{0, 1\}$. By definition, $\mathbf{w}_i \in \mathcal{A}_i$. □

To establish that $\mathcal{A}_\alpha \subseteq \mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha)$, we need results on the geometry of sets generated by normalized matrices that sum to $\mathbf{I}$ (see Lemma 8 in the Appendix). This is not true for generic scale parameters (see Lemma 9), but it holds under the restrictions of Theorem 3.

Lemma 7: Consider $\mathcal{A}_0$ and $\mathcal{A}_1$ as defined in (8). If,

1) $\mathbf{\Psi}_1 = \mathbf{\Psi}_0$, or,
2) $\mathbf{\Psi}_i = \psi_i\mathbf{I}$ with $\psi_i > 0$ for $i \in \{0, 1\}$, or,
3) $\mathbf{\Psi}_i = \mathbf{\Lambda}_i$ with $\mathbf{\Lambda}_i \in \mathcal{D}_n^+$ for $i \in \{0, 1\}$, with $\mathrm{Q}^{-1}(z) = \prod_{j=1}^n [a_j(z), b_j(z)]$ for $a_j(z), b_j(z) \in \overline{\mathbb{R}}$ and for all $j \in \{1, 2, \ldots, n\}$ and any $z \in \mathcal{Z}$,

then $\mathcal{A}_\alpha \subseteq \mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha)$. Here, $\mathcal{D}_n^+$ is the space of $n \times n$ non-negative diagonal matrices.

Proof: Let $\alpha_0 = 1 - \alpha$ and $\alpha_1 = \alpha$ and consider the matrices

$$\mathbf{C}_i = (\alpha_0\mathbf{\Psi}_0 + \alpha_1\mathbf{\Psi}_1)^{-1}\alpha_i\mathbf{\Psi}_i$$

for $i \in \{0, 1\}$. Consider also that $\mathbf{C}_0 + \mathbf{C}_1 = \mathbf{I}$.

Let $\mathbf{w} \in \mathcal{A}_\alpha$. Then, there are $\mathbf{w}_i \in \mathcal{A}_i$ for $i \in \{0, 1\}$ such that $\mathbf{w} = \alpha_0\mathbf{w}_0 + \alpha_1\mathbf{w}_1$. Furthermore, by (8) we have that $\mathbf{w}_i = \mathbf{\Psi}_i\mathbf{y}_i - \mathbf{S}\mathbf{x}_i$ for $i \in \{0, 1\}$, where $\mathbf{y}_i \in \mathrm{Q}^{-1}(z)$. Therefore,

$$\begin{aligned}\mathbf{w} &= \sum_{i=0}^{1}\left(\alpha_i\mathbf{\Psi}_i\mathbf{y}_i - \mathbf{S}\alpha_i\mathbf{x}_i\right) \\ &= \left(\alpha_0\mathbf{\Psi}_0 + \alpha_1\mathbf{\Psi}_1\right)\left(\mathbf{C}_0\mathbf{y}_0 + \mathbf{C}_1\mathbf{y}_1\right) - \mathbf{S}\left(\alpha_0\mathbf{x}_0 + \alpha_1\mathbf{x}_1\right).\end{aligned}$$

Then, $\mathbf{w} \in \mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha)$ if and only if $\mathbf{y} = \mathbf{C}_0\mathbf{y}_0 + \mathbf{C}_1\mathbf{y}_1 \in \mathrm{Q}^{-1}(z)$.

If condition a) is fulfilled, then $\mathbf{C}_i = \alpha_i\mathbf{I}$ and $\mathbf{y} = \mathbf{y}_\alpha$. Because $\mathrm{Q}^{-1}(z)$ is convex, $\mathbf{y} \in \mathrm{Q}^{-1}(z)$. If condition b) is fulfilled, then $\mathbf{C}_i = \tilde{\alpha}_i\mathbf{I}$ with $\tilde{\alpha}_i = \alpha_i\psi_i/(\alpha_0\psi_0 + \alpha_1\psi_1)$, and $\mathbf{y}$ is a convex combination of $\mathbf{y}_0$ and $\mathbf{y}_1$, i.e., $\mathbf{y} = \mathbf{y}_{\tilde{\alpha}_1}$. Because $\mathrm{Q}^{-1}(z)$ is convex, $\mathbf{y} \in \mathrm{Q}^{-1}(z)$. If condition c) is fulfilled, then the $\mathbf{C}_i \in \mathcal{D}_n([0, 1])$ for $i \in \{0, 1\}$, where $\mathcal{D}_n([0, 1])$ is the set of $n \times n$ diagonal matrices with elements in [0,1]. By Lemma 8, we then have $\mathbf{y} \in \prod_{j=1}^n [\mathbf{y}_1[j], \mathbf{y}_2[j]]$. Because $\mathrm{Q}^{-1}(z) = \prod_{j=1}^n [a_j(z), b_j(z)]$, and $a_j(z) \leq \mathbf{y}_1[j], \mathbf{y}_2[j] \leq b_j(z)$, we have

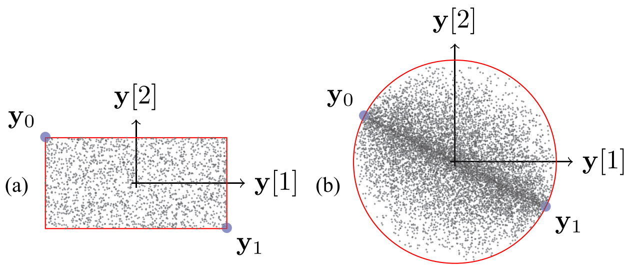

Fig. 3. In black, points obtained by combinations with matrices whose sum is the identity matrix, i.e., $\mathbf{C}\mathbf{y}_1 + (\mathbf{I} - \mathbf{C})\mathbf{y}_0$, where $\mathbf{C}$ were, in a), random diagonal matrices from $\mathcal{D}_n([0,1])$, and, in b), random positive semidefinite matrices from $\mathcal{M}_n^+(\mathbb{R})$ with $\rho(\mathbf{C}) \leq 1$. In blue, $\mathbf{y}_0$ and $\mathbf{y}_1$.

that $\mathbf{y} \in \mathrm{Q}^{-1}(z)$. Therefore, if either a), b) or c) are given, $\mathbf{w} \in \mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha)$.

We can now combine Lemmas 5–7 with Theorem 4 to show Theorem 3.

Proof of Theorem 3: Consider $\mathcal{A}_0$ and $\mathcal{A}_1$ as defined in (8). By Lemma 5, they are convex sets. By Lemmas 6 and 7, if a), b) or c) are fulfilled, we have that $\mathcal{A}_\alpha = \mathcal{W}_z(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha)$. Then, Theorem 4 and (6) yield

$$\begin{aligned}\mathcal{L}(\mathbf{x}_\alpha, \mathbf{\Psi}_\alpha; z) &= \mathrm{P}_{\mathbf{w}}\left[\mathcal{A}_\alpha\right] \\ &\geq \mathrm{P}_{\mathbf{w}}\left[\mathcal{A}_1\right]^\alpha \mathrm{P}_{\mathbf{w}}\left[\mathcal{A}_0\right]^{(1-\alpha)} \\ &= \mathcal{L}(\mathbf{x}_1, \mathbf{\Psi}_1; z)^\alpha\, \mathcal{L}(\mathbf{x}_0, \mathbf{\Psi}_0; z)^{1-\alpha}.\end{aligned}$$

□

Acknowledgment

We acknowledge access to the facilities and expertise of the CIBM Center for Biomedical Imaging, a Swiss research center of excellence founded and supported by Lausanne University Hospital (CHUV), University of Lausanne (UNIL), École polytechnique fédérale de Lausanne (EPFL), University of Geneva (UNIGE), and Geneva University Hospitals (HUG).

Appendix
Matrix Combinations

In the proof of Lemma 7, we use that sets of the form $\prod_{j=1}^n [a_j, b_j]$ with $a_j, b_j \in \overline{\mathbb{R}}$ are closed with respect to the generalization of convex combinations to diagonal matrices. In Fig. 3 a, we include an illustration of a practical case in 2D. Here, we provide the proof. To our knowledge, this has not been reported before.

Lemma 8 (Diagonal Matrices Whose Sum is the Identity Generate Squares): Let $\mathcal{D}_n([0,1])$ be the set of $n \times n$ diagonal matrices with elements in [0,1], and let $\mathbf{y}_0, \mathbf{y}_1 \in \mathbb{R}^n$. Then,

$$\begin{aligned}\mathcal{H} &:= \{\mathbf{C}\mathbf{y}_0 + (\mathbf{I} - \mathbf{C})\mathbf{y}_1 : \mathbf{C} \in \mathcal{D}_n\,[0,1]\} \\ &= \prod_{j=1}^n [\mathbf{y}_0[j], \mathbf{y}_1[j]] =: \mathcal{H}_\blacksquare\,.\end{aligned}$$

Proof: For $\mathcal{H}_\blacksquare \subseteq \mathcal{H}$, let $\mathbf{y} \in \mathcal{H}_\blacksquare$. If $\alpha_j = (\mathbf{y}[j] - \mathbf{y}_1[j])/(\mathbf{y}_0[j] - \mathbf{y}_1[j])$, then $\alpha_j \in [0,1]$. If $\mathbf{C} \in \mathcal{D}_n([0,1])$ is the diagonal matrix such that $\mathbf{C}[j,j] = \alpha_j$, then $\mathbf{C}\mathbf{y}_0 + (\mathbf{I} - \mathbf{C})\mathbf{y}_1 = \mathbf{y}$. Thus, $\mathbf{y} \in \mathcal{H}$.

For $\mathcal{H} \subseteq \mathcal{H}_\blacksquare$, let $\mathbf{y} \in \mathcal{H}$. Then, we have that $\alpha_j = \mathbf{C}[j,j] \in [0,1]$, and $\mathbf{y}[j] = \alpha_j \mathbf{y}_1[j] + (1 - \alpha_j)\mathbf{y}_0[j]$, and thus, $\mathbf{y}[j] \in [\mathbf{y}_0[j], \mathbf{y}_1[j]]$. Therefore, $\mathbf{y} \in \mathcal{H}_\blacksquare$.

On the one hand, Lemma 8 allows for the most general result in terms of the scale parameter $\mathbf{\Psi}$ we have obtained (Theorem 3.c). On the other hand, the corresponding result for arbitrary scale parameters suggests that the strategy behind our proof of Theorem 3 might not generalize well. We include it here in Lemma 9 for completeness. To our knowledge, it has not been reported before.

Lemma 9 (Positive Semidefinite Matrices that Sum to the Identity Generate Balls[1]): Let $\mathcal{M}_n^\oplus(\mathbb{R})$ be the set of real $n \times n$ symmetric positive-semidefinite matrices with spectral radius less or equal than 1, i.e., $\rho(\mathbf{C}) \leq 1$, and $\mathbf{y}_0, \mathbf{y}_1 \in \mathbb{R}^n$. Then

$$\begin{aligned}\mathcal{S} &:= \left\{\mathbf{C}\mathbf{y}_0 + (\mathbf{I} - \mathbf{C})\,\mathbf{y}_1 : \mathbf{C} \in \mathcal{M}_n^\oplus\,(\mathbb{R})\right\} \\ &= \mathcal{B}\left(\frac{\mathbf{y}_0 + \mathbf{y}_1}{2}, \frac{1}{2}\left\|\mathbf{y}_1 - \mathbf{y}_0\right\|_2\right) =: \mathcal{S}_\bullet\,,\end{aligned}$$

where $\mathcal{B}(\mathbf{y}_c, r)$ is the closed ℓ_2 ball centered at $\mathbf{y}_c \in \mathbb{R}^n$ with radius $r > 0$.

Proof: For $\mathcal{S}_\bullet \subseteq \mathcal{S}$, let $\mathbf{y} \in \mathcal{S}$. Then, there is a $\mathbf{C} \in \mathcal{M}_n^\oplus(\mathbb{R})$ such that $\mathbf{y} = \mathbf{C}\mathbf{y}_0 + (\mathbf{I} - \mathbf{C})\mathbf{y}_1$. If $\mathbf{y}_c = (\mathbf{y}_0 + \mathbf{y}_1)/2$, then

$$\begin{aligned}\left\|\mathbf{y} - \mathbf{y}_c\right\|_2 &= \left\|\left(\mathbf{C} - \frac{\mathbf{I}}{2}\right)(\mathbf{y}_0 - \mathbf{y}_1)\right\|_2 \\ &\leq \left\|\mathbf{C} - \frac{\mathbf{I}}{2}\right\|_{2\to 2} \left\|\mathbf{y}_0 - \mathbf{y}_1\right\|_2 \\ &\leq \frac{1}{2}\left\|\mathbf{y}_1 - \mathbf{y}_0\right\|_2\,.\end{aligned}$$

Therefore, $\mathbf{y} \in \mathcal{S}_\bullet$. Here, $\|\cdot\|_{2\to 2}$ is a the operator norm of a matrix with respect to the ℓ_2 norm, and we have used that it coincides with the spectral radius $\rho(\cdot)$ for Hermitian matrices.

For $\mathcal{S}_\bullet \subseteq \mathcal{S}$, let $\mathbf{y} \in \mathcal{S}_\bullet$ and $\mathbf{y}_c = (\mathbf{y}_0 + \mathbf{y}_1)/2$. Then, consider $\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{y}_1$ and $\tilde{\mathbf{y}}_0 = \mathbf{y}_0 - \mathbf{y}_1$. Because $\mathbf{y} \in \mathcal{S}_\bullet$, we have that $\|\tilde{\mathbf{y}} - \tilde{\mathbf{y}}_0/2\|_2^2 = \|\mathbf{y} - \mathbf{y}_c\|_2^2 \leq (\|\mathbf{y}_1 - \mathbf{y}_0\|_2/2)^2 = (\|\tilde{\mathbf{y}}_0\|_2/2)^2$. Expanding the squares, we obtain $0 \leq \|\tilde{\mathbf{y}}\|_2^2 \leq \tilde{\mathbf{y}}_0^{\mathrm{T}}\tilde{\mathbf{y}}$. If we consider then the matrix $\mathbf{C} = \tilde{\mathbf{y}}\tilde{\mathbf{y}}^{\mathrm{T}}/\tilde{\mathbf{y}}^{\mathrm{T}}\tilde{\mathbf{y}}_0$, we see that it is a rank 1 matrix with a single non-zero eigenvalue $\lambda = \mathrm{tr}(\mathbf{C}) = \|\tilde{\mathbf{y}}\|^2/\tilde{\mathbf{y}}^{\mathrm{T}}\tilde{\mathbf{y}}_0 \in [0,1]$, i.e., $\mathbf{C} \in \mathcal{M}_n^\oplus(\mathbb{R})$. Furthermore, $\mathbf{C}(\mathbf{y}_0 - \mathbf{y}_1) = \mathbf{C}\tilde{\mathbf{y}}_0 = \tilde{\mathbf{y}} = \mathbf{y} - \mathbf{y}_1$, i.e., $\mathbf{y} = \mathbf{C}\mathbf{y}_0 + (\mathbf{I} - \mathbf{C})\mathbf{y}_1$. Therefore, $\mathbf{y} \in \mathcal{S}$. □

Remark 10: To apply our proof technique for Lemma 7 to arbitrary scale parameters $\mathbf{\Psi} \in \mathcal{M}_n^+(\mathbb{R})$, we would need to restrict the quantization regions so that $\mathbf{y}_0, \mathbf{y}_1 \in \mathrm{Q}^{-1}(z)$ implies $\mathcal{S}_\bullet \subset \mathrm{Q}^{-1}(z)$. Our intuitive understanding is that only trivial quantizers would fulfill this property. However, because Theorem 4 is sufficient but not necessary, Lemma 9 does not preclude likelihood logconcavity results for that case.

[1]See www.geogebra.org/m/hdxtmz3b and www.geogebra.org/m/tskjev2m for 2D and 3D dynamic examples of Lemma 9 based on the SVD of any positive semidefinite matrix.

REFERENCES

[1] B. Widrow and I. Kollár, *Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control, and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[2] A. Azizzadeh, R. Mohammadkhani, S. V. A.-D. Makki, and E. Björnson, "BER performance analysis of coarsely quantized uplink massive MIMO," *Signal Process.*, vol. 161, pp. 259–267, 2019.

[3] S. Li, X. Li, X. Wang, and J. Liu, "Decentralized sequential composite hypothesis test based on one-bit communication," *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3405–3424, Jun. 2017.

[4] J. Ren, T. Zhang, J. Li, and P. Stoica, "Sinusoidal parameter estimation from signed measurements via majorization-minimization based RELAX," *IEEE Trans. Signal Process.*, vol. 67, no. 8, pp. 2173–2186, Apr. 2019.

[5] S. Khobahi, N. Naimipour, M. Soltanalian, and Y. C. Eldar, "Deep signal recovery with one-bit quantization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 2987–2991.

[6] C. K. Wen, C. J. Wang, S. Jin, K. K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2541–2556, May 2016.

[7] H. Zayyani, M. Korki, and F. Marvasti, "Dictionary learning for blind one bit compressed sensing," *IEEE Signal Process. Lett.*, vol. 23, no. 2, pp. 187–191, Feb. 2016.

[8] H. Zayyani, M. Korki, and F. Marvasti, "A distributed 1-bit compressed sensing algorithm robust to impulsive noise," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1132–1135, Jun. 2016.

[9] M. S. Stein, S. Bar, J. A. Nossek, and J. Tabrikian, "Performance analysis for channel estimation with 1-bit ADC and unknown quantization threshold," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2557–2571, May 2018.

[10] C. C. Aggarwal and P. S. Yu, "A general survey of privacy-preserving data mining models and algorithms," in *Proc. Privacy-Preserving Data Mining*, 2008, pp. 11–52.

[11] P. Gao, R. Wang, M. Wang, and J. H. Chow, "Low-rank matrix recovery from noisy, quantized, and erroneous measurements," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2918–2932, Jun. 2018.

[12] Y. Zhang, C. Liao, M. Liu, S. Lu, and Z. Li, " Optimization of 3D shell electrode detectors-A type of honeycomb shell electrode detector," *AIP Adv.*, vol. 8, no. 12, 2018, Art. no. 125309.

[13] M. Danielsson, M. Persson, and M. Sjölin, "Photon-counting x-ray detectors for CT," *Phys. Med. Biol.*, vol. 66, no. 3, 2021, Art. no. 03TR01.

[14] G. Köklü, J. Ghaye, R. Beuchat, G. De Micheli, Y. Leblebici, and S. Carrara, "Quantitative comparison of commercial CCD and custom-designed CMOS camera for biological applications," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2012, pp. 2063–2066.

[15] J. W. Pratt, "Concavity of the log likelihood," *J. Amer. Stat. Assoc.*, vol. 76, no. 373, pp. 103–106, 1981.

[16] J. Burridge, "Some unimodality properties of likelihoods derived from grouped data," *Biometrika*, vol. 69, no. 1, pp. 145–151, 1982.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[18] A. Conti, D. Panchenko, S. Sidenko, and V. Tralli, "Log-concavity property of the error probability with application to local bounds for wireless communications," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2766–2775, Jun. 2009.

[19] E. J. Msechu and G. B. Giannakis, "Sensor-centric data reduction for estimation with WSNs via censoring and quantization," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 400–414, Jan. 2012.

[20] A. Prékopa, "Logarithmic concave measures and functions," *Acta Scientiarum Mathematicarum*, vol. 34, no. 1, pp. 334–343, 1973.