High-Noise Asymptotics of the Ziv-Zakai Bound

Alex Dytso†, Martina Cardone*, Ian Zieder†
† New Jersey Institute of Technology, Newark, NJ 07102, USA, Email: {alex.dytso,ihz3}@njit.edu
* University of Minnesota, Minneapolis, MN 55404, USA, Email: cardo089@umn.edu

***Abstract*—The Ziv-Zakai bound is a well-known lower bound on the minimum mean squared error. This letter analyzes the performance of this bound in the practically relevant high-noise regime for a broad family of observation models. The goal is to understand whether this bound is tight, and in which scenarios it should be used. It is shown that, while the Ziv-Zakai bound is tight for a certain class of symmetric distributions, in general, it is *not* tight in the high-noise regime.**

I. Introduction

A classical Bayesian estimation problem seeks to estimate a realization of a random variable X from the noisy observation $\mathbf{Y}$[1]. In this work, we assume that X is a scalar random variable[2], and $\mathbf{Y}$ can be an arbitrary random object (i.e., random vector or random process). The minimum mean squared error (MMSE) is defined and denoted by

$$\mathrm{mmse}(X|\mathbf{Y}) = \mathbb{E}\left[(X - \mathbb{E}[X|\mathbf{Y}])^2\right], \tag{1}$$

where $\mathbb{E}[\cdot]$ denotes the expected value. The MMSE is by far the most popular fidelity criterion for assessing the quality of estimating X from $\mathbf{Y}$. However, the MMSE is often difficult to compute and one needs to rely on bounds. The focus usually falls on the lower bounds as these are typically much more difficult to derive than the upper bounds. The latter are usually derived by considering suboptimal estimators, e.g., linear estimators [3], [4] or stochastic estimators [5].

Several different families of lower bounds are available in the literature. For example, the Weiss–Weinstein family [6] is a popular family of bounds, and it includes important bounds such as the Bayesian Cramér-Rao bound [7] (also known as the Van Trees bound), the Bobrovsky–Zakai bound [8], a Barankin-type bound [9], and the Bobrovsky-Mayer-Wolf-Zakai bound [10]. Due to space limitations, we do not attempt to survey the literature on this vast subject and refer the reader to [1], [11], [12], [13] for a comprehensive literature review.

An interesting bound that beautifully connects binary hypothesis testing and estimation is known as the Ziv-Zakai bound [1], [14], [15], [16], [17], [18]. This bound is believed to be one of the tightest bounds available in the literature. Before presenting the bound we need the following two notions.

The work of M. Cardone was supported in part by the U.S. National Science Foundation under Grant CCF-1849757.

[1]As with any Bayesian setting, we assume that the joint distribution of $(X, \mathbf{Y})$ is known, at least in principle.

[2]Our focus is on analyzing the Ziv-Zakai bound for which, to the best of our knowledge, a vector version is not known. The version that has been derived in [1] (and its useful application in bearing estimation [2]), in fact, does not directly bound the MMSE, but a quadratic expression involving the MMSE matrix (see [1, Property 6]). Thus, we here assume that X is scalar.

Definition 1. *The* valley-filling *function acting on a function $f : \mathbb{R} \to \mathbb{R}$ is defined as*

$$\mathcal{V}\{f(x)\} = \sup_{\varepsilon \geq 0} f(x+\varepsilon), \ x \in \mathbb{R}. \tag{2}$$

Definition 2. *For a given $x_0, x_1 \in \mathbb{R}$, $P_e[x_0, x_1, p_0, p_1]$ denotes* the minimum probability of error *(obtained by using the optimal likelihood ratio test) for the following binary hypothesis testing problem,*

$$H_0 : \mathbf{Y} \sim P_{\mathbf{Y}|X}(\mathbf{y}|x_0),$$
$$H_1 : \mathbf{Y} \sim P_{\mathbf{Y}|X}(\mathbf{y}|x_1),$$

where

$$\mathsf{Pr}(H_0) = p_0, \ \mathsf{Pr}(H_1) = 1 - \mathsf{Pr}(H_0) = p_1.$$

The general Ziv-Zakai lower bound is stated next [1], [16].

Theorem 1. (Ziv-Zakai Lower Bound.) *Consider a pair of random variables $(X, \mathbf{Y})$ where X has probability density function (pdf) $f_X(x)$ and where the noisy observation model $\mathbf{Y}|X = x$ is governed by the distribution $P_{\mathbf{Y}|X}(\mathbf{y}|x)$. Then, we have*

$$\mathrm{mmse}(X|\mathbf{Y}) \geq \overline{\mathsf{LB}_{\mathsf{ZZ}}}(X|\mathbf{Y}), \tag{3}$$

where

$$\overline{\mathsf{LB}_{\mathsf{ZZ}}}(X|\mathbf{Y}) = \frac{1}{2}\int_0^\infty \mathcal{V}\left\{\int_{-\infty}^{\infty} P_e[x, x+h, p_0(x,h), p_1(x,h)] \cdot (f_X(x) + f_X(x+h)) \, \mathrm{d}x\right\} h \, \mathrm{d}h, \tag{4}$$

with

$$p_0(x,h) = \frac{f_X(x)}{f_X(x) + f_X(x+h)}, \ p_1(x,h) = 1 - p_0(x,h).$$

The valley-filling function in Theorem 1 introduces an extra layer of optimization which can make the bound difficult to evaluate. Thus, one often considers a loosened version of the Ziv-Zakai bound that drops the valley-filling function, that is,

$$\mathsf{LB}_{\mathsf{ZZ}}(X|\mathbf{Y}) = \frac{1}{2}\int_0^\infty \int_{-\infty}^{\infty} P_e[x, x+h, p_0(x,h), p_1(x,h)] \cdot (f_X(x) + f_X(x+h))\, h \, \mathrm{d}x \, \mathrm{d}h. \tag{5}$$

In this work, our main goal is to understand the behavior of the Ziv-Zakai bounds in (4) and in (5) in the practically relevant high-noise regime. Different from the low-noise regime where several lower bounds are known to perform well [13], [19], in the high-noise the same is not true in general. Thus, it is of interest to understand if the Ziv-Zakai bound is tight in high-noise. The analysis of this regime is an important benchmark for the performance, especially in wireless scenarios, where high-noise represents a weak signal scenario, and has received some attention in various contexts [3], [20], [21], [22].

Notation. Random variables are denoted by upper case letters and their instances by lower case letters. The expected value of a random variable X and its variance are denoted by $\mathbb{E}[X]$ and $\mathsf{Var}(X)$, respectively.

A. Noise Models Under Consideration

We here describe a family of noise distributions for which our results hold. First, in order to take limits and quantify the strength of the noise, we need to be able to parameterize our model and hence, we make the following assumption,

A1: $P_{\mathbf{Y}|X}$ can be parameterized in terms of the parameter $\eta \ge 0$, i.e., $P_{\mathbf{Y}|X}(\mathbf{y}|x) = P_{\mathbf{Y}|X}(\mathbf{y}|x;\eta)$ for all $(x,\mathbf{y})$. We refer to the parameter η as the *noise level*.

Second, we require that the performance of our system degrades as the noise level increases. Towards this end, we make the following assumption,

A2: For the sequence of noisy observation models $\{P_{\mathbf{Y}|X}(\cdot|\cdot;\eta)\}_{\eta\ge 0}$, we parameterize $P_e\left[x_0,x_1,p_0,p_1\right]= P_e\left[\eta;x_0,x_1,p_0,p_1\right]$, and we assume that the following holds. For every (x_0,x_1,p_0,p_1) where $x_0 \neq x_1$:

A2a: $\eta \mapsto P_e[\eta;x_0,x_1,p_0,p_1]$ is non-decreasing; and

A2b: $$\lim_{\eta\to\infty} P_e[\eta;x_0,x_1,p_0,p_1] = \min\{p_0,p_1\}. \tag{6}$$

All of the assumptions above are rather natural. In particular, assumption **A2a** simply states that the probability of error for binary detection can not decrease as the noise level increases. Assumption **A2b** states that when the observation is completely dominated by the noise, the best strategy is to guess the $x_i, i \in \{0,1\}$ with the largest probability.

Most of the observation models encountered in practice satisfy the above assumptions. We now give a few examples.

- *Additive White Gaussian Model:* Let
$$Y = X + \sqrt{\eta} Z, \tag{7}$$
with X and Z being independent, and Z being a standard Gaussian random variable. In this case, the noise level parameter $\eta > 0$ is known as the *noise power* [4].
- *Poisson Noise Model:* For $x \ge 0$, $y \in \mathbb{N}_0$, let
$$P_{Y|X}(y|x) = \frac{(x+\eta)^y \mathrm{e}^{-(x+\eta)}}{y!}; \tag{8}$$
in this case, the noise level parameter $\eta > 0$ is known as the *dark current* parameter [23], [24].
- *Binary Symmetric Model:* For $x \in \{0,1\}$ and $y \in \{0,1\}$, let
$$P_{Y|X}(y|x) = \begin{cases} 1-\eta & x = y, \\ \eta & x \neq y, \end{cases} \tag{9}$$
where the noise level parameter $\eta \in \left(0, \frac{1}{2}\right)$ is known as the *cross over probability* [25].

See also [26] for an example on how to define the noise level parameter η in the context of the exponential family.

For the models that satisfy all of the above assumptions, we parameterize $\mathrm{mmse}(X|\mathbf{Y})$ as a function of X and η. Thus, in what follows we denote it as $\mathrm{mmse}(X,\eta)$. Similarly, we use $\overline{\mathsf{LB}_{\mathsf{ZZ}}}(X,\eta)$ and $\mathsf{LB}_{\mathsf{ZZ}}(X,\eta)$ for $\overline{\mathsf{LB}_{\mathsf{ZZ}}}(X|\mathbf{Y})$ in (4) and $\mathsf{LB}_{\mathsf{ZZ}}(X|\mathbf{Y})$ in (5), respectively.

B. Goals and Contributions

Note that, in general, for every observation model, by using $\mathbb{E}[X]$ as a suboptimal estimator of X, we always have that
$$\mathrm{mmse}(X|\mathbf{Y}) \le \mathsf{Var}(X). \tag{10}$$
The variance of X is, in fact, the MMSE of estimating X without any knowledge of $\mathbf{Y}$ (i.e., the error is due to the blind guess). For practically relevant observation models, in the high-noise regime, we expect that the output $\mathbf{Y}$ carries no information. Thus, it is reasonable to assume that the following limit holds,
$$\lim_{\eta\to\infty} \mathrm{mmse}(X,\eta) = \mathsf{Var}(X). \tag{11}$$
The interested reader is referred to [3] and [23] where the above limit has been formally established for the Gaussian noise model and the Poisson noise model, respectively.

The upper bound in (10) and the asymptotic in (11) can be used to assess whether a given lower bound on the MMSE is tight or not. For example, consider the ubiquitous Bayesian Cramér-Rao (CR) bound [7], which for the Gaussian noise case in (7) is given by
$$\mathsf{LB}_{\mathsf{CR}}(X,\eta) = \frac{1}{J(X) + \frac{1}{\eta}}, \tag{12}$$
where $J(X) = \mathbb{E}\left[\left(\frac{f_X'(X)}{f_X(X)}\right)^2\right]$ is the Fisher information. The CR bound holds provided that the pdf $f_X(x)$ is equal to zero on the boundaries of the support of X [27]. For example, it does not hold for an exponential distribution, i.e., $f_X(x) = \lambda \mathrm{e}^{-\lambda x} 1_{\{x\ge 0\}}$, since $f_X(0^+) = \lambda$. It is not difficult to see that
$$\lim_{\eta\to\infty} \mathsf{LB}_{\mathsf{CR}}(X,\eta) = \frac{1}{J(X)}. \tag{13}$$
Consequently, from (11) and (13) we have that the CR bound is tight for large η if $\frac{1}{J(X)} = \mathsf{Var}(X)$. However, note that $\frac{1}{J(X)}$ is equal to $\mathsf{Var}(X)$ only when X is Gaussian [27], [28]. Thus, from (13) we conclude that, in general, the CR bound is not tight in the high-noise regime.

The goal of this paper is to perform a similar analysis for both of the Ziv-Zakai bounds in (4) and (5) and find the high-noise asymptotics for the broad family of observation models that satisfy assumptions **A1** and **A2**. Specifically (provided that the limits exist) we are interested in characterizing the following limits,
$$\lim_{\eta\to\infty} \mathsf{LB}_{\mathsf{ZZ}}(X,\eta) = \mathbb{V}(X), \tag{14a}$$
$$\lim_{\eta\to\infty} \overline{\mathsf{LB}_{\mathsf{ZZ}}}(X,\eta) = \overline{\mathbb{V}}(X). \tag{14b}$$
From (4), (5) and (10) we have the following inequalities,
$$\mathbb{V}(X) \le \overline{\mathbb{V}}(X) \le \mathsf{Var}(X). \tag{15}$$

Our goal is to establish whether these inequalities hold with equality. To this end, in Section II-A we first characterize $\mathbb{V}(X)$ and $\overline{\mathbb{V}}(X)$ in (14). Then, in Section II-B, we provide examples under which either all or a subset of the inequalities in (15) hold and/or do not hold. This implies that there are examples of distributions on X for which the Ziv-Zakai bound is tight, and there are examples for which it is not tight. Finally, in Section II-C we compare the Ziv-Zakai and the CR bounds.

II. Main Results

A. High-Noise Asymptotics

The next theorem (proof in Appendix A) provides the behavior of the two Ziv-Zakai lower bounds in (14).

Theorem 2. *Under the assumptions* **A1** *and* **A2**, *we have the following,*

$$\mathbb{V}(X) = \frac{1}{4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \min\{f_X(y), f_X(x)\}\ |y-x|\ \mathrm{d}x\ \mathrm{d}y,$$

$$\overline{\mathbb{V}}(X) = \frac{1}{4}\int_{-\infty}^{\infty} \mathcal{V}\Big\{ \int_{-\infty}^{\infty} \min\{f_X(x), f_X(x+h)\}\,\mathrm{d}x \Big\}\ |h|\ \mathrm{d}h.$$

B. Examples

The examples below (detailed computations in Appendix B) show that, while there are cases for which even the loosened version of the Ziv-Zakai bound is tight (i.e., $\mathbb{V}(X) = \mathsf{Var}(X)$), in general, neither of the bounds in (14) is tight.

Example 1. Suppose that $f_X(x) = g(|x|)$ where $g : [0,\infty) \to [0,\infty)$ is non-increasing. Then,

$$\mathbb{V}(X) = \overline{\mathbb{V}}(X) = \mathsf{Var}(X).$$

Thus, both versions of the Ziv-Zakai bound agree and are tight, i.e., the valley-filling function is not needed in this case. This example encompasses a broad range of widely used symmetric distributions (e.g., Gaussian, Laplace, generalized normal). $\square$

Example 2. Suppose that $f_X(x)$ is non-increasing on (a,b) where $-\infty < a < b \le \infty$, and zero elsewhere. Then,

$$\mathbb{V}(X) = \overline{\mathbb{V}}(X) = \frac{\mathsf{Var}(X) + (a - \mathbb{E}[X])^2}{4}.$$

Thus, the two versions of the Ziv-Zakai bound agree, but are *not* tight. For instance, assume that X has an exponential pdf with parameter λ; we have that $\mathsf{Var}(X) = 1/\lambda^2$ and $\mathbb{V}(X) = \overline{\mathbb{V}}(X) = 1/(2\lambda^2)$, i.e., the Ziv-Zakai bound is off by a factor of two and hence, it can be substantially suboptimal. $\square$

Example 3. Suppose that $0 < a < b$ and

$$f_X(x) = \frac{1}{2(b-a)}\left(\mathrm{rect}(x; -b, -a) + \mathrm{rect}(x; a, b)\right),$$

where $x \mapsto \mathrm{rect}(x; a, b)$ is the unit-height rectangle with support over the interval (a,b). In other words, the distribution of X is a mixture of two uniform distributions. Then, for $0 < a < b < 3a$, we have that

$$\mathsf{Var}(X) = \frac{a^2 + ab + b^2}{3},$$

and

$$\mathbb{V}(X) = \mathsf{Var}(X) - \frac{a(a+b)}{2},$$

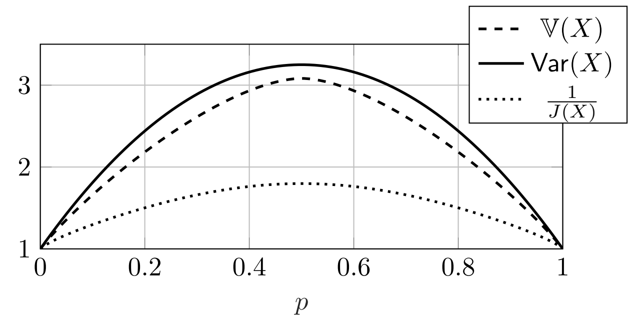

Fig. 1: Mixed Gaussian $f_X = p\mathcal{N}(-1,1) + (1-p)\mathcal{N}(2,1)$ where $p \in [0,1]$.

$$\overline{\mathbb{V}}(X) = \mathbb{V}(X) + \frac{7a^2 + 10ab - b^2}{32}$$
$$= \mathsf{Var}(X) - \frac{(3a+b)^2}{32}.$$

Thus, the Ziv-Zakai bound with the valley-filling function can be strictly better than the one without it, yet not optimal. $\square$

We were not able to identify an example for which $\mathbb{V}(X) < \overline{\mathbb{V}}(X) = \mathsf{Var}(X)$, i.e., a case for which the Ziv-Zakai bound with a valley-filling function is optimal, but the bound without the valley-filling function is strictly sub-optimal.

C. Comparison with the CR Bound

An interesting question that arises is: Does the Ziv-Zakai bound outperform the CR bound[3]? To show this analytically, one would need to demonstrate the following inequality (inspired by the limit in (13)),

$$\frac{1}{J(X)} \le \overline{\mathbb{V}}(X).$$

In Fig. 1, this inequality is numerically verified for a mixed Gaussian distribution.

Appendix A
Proof of Theorem 2

Before proceeding with the proof, we will need the following facts. First, by using (2), note that if $f(x) \le g(x)$ for all x, then for all x it holds that

$$\mathcal{V}\{f(x)\} \le \mathcal{V}\{g(x)\}. \tag{16}$$

Second, the valley-filling function is lower semicontinuous, i.e., for any sequence of functions $\{f_n\}_{n=1}^{\infty}$, we have that

$$\liminf_{n\to\infty} \mathcal{V}\{f_n(x)\} \ge \mathcal{V}\{\liminf_{n\to\infty} f_n(x)\}. \tag{17}$$

To see this recall that $\liminf_{n\to\infty} f_n(x) = \sup_{n\ge 0}\inf_{m\ge n} f_m(x)$ and note that

$$\begin{aligned}\liminf_{n\to\infty} \mathcal{V}\{f_n(x)\} &= \sup_{n\ge 0}\inf_{m\ge n}\sup_{\varepsilon\ge 0} f_m(x+\varepsilon)\\ &\ge \sup_{\varepsilon\ge 0}\sup_{n\ge 0}\inf_{m\ge n} f_m(x+\varepsilon)\\ &= \mathcal{V}\{\liminf_{n\to\infty} f_n(x)\},\end{aligned} \tag{18}$$

[3]We note that the Ziv-Zakai bound holds for a larger family of distributions than the CR bound as the pdf does not need to be differentiable or even continuous.

where the inequality follows from the max-min inequality.

We now consider the behavior of $\overline{\mathsf{LB}_{\mathsf{ZZ}}}(X,\eta)$. We have

$$\begin{aligned}
&2\overline{\mathsf{LB}_{\mathsf{ZZ}}}(X,\eta)\\
&= \int_0^\infty \mathcal{V}\left\{\int_{-\infty}^\infty P_e\left[\eta; x, x+h, p_0(x,h), p_1(x,h)\right]\right.\\
&\qquad \left. \cdot \left(f_X(x)+f_X(x+h)\right) \mathrm{d}x\right\} h \ \mathrm{d}h\\
&\overset{(a)}{\le} \int_0^\infty \mathcal{V}\left\{\int_{-\infty}^\infty P_e\left[\infty; x, x+h, p_0(x,h), p_1(x,h)\right]\right.\\
&\qquad \left. \cdot \left(f_X(x)+f_X(x+h)\right) \mathrm{d}x\right\} h \ \mathrm{d}h\\
&\overset{(b)}{=} \int_0^\infty \mathcal{V}\Big\{\int_{-\infty}^\infty \min\left\{f_X(x), f_X(x+h)\right\} \mathrm{d}x\Big\} \ h \ \mathrm{d}h, \quad (19)
\end{aligned}$$

where (a) follows by using (16) and by noting that

$$P_e\left[\eta; x, x+h, p_0, p_1\right] \le P_e\left[\infty; x, x+h, p_0, p_1\right],$$

which is a consequence of the assumption **A2a**; and (b) follows from the assumption **A2b**.

Next, we note that

$$\begin{aligned}
&\liminf_{\eta\to\infty} \ 2\overline{\mathsf{LB}_{\mathsf{ZZ}}}(X,\eta)\\
&\overset{(c)}{\ge} \int_0^\infty \liminf_{\eta\to\infty} \mathcal{V}\left\{\int_{-\infty}^\infty P_e\left[\eta; x, x+h, p_0(x,h), p_1(x,h)\right]\right.\\
&\qquad \left. \cdot \left(f_X(x)+f_X(x+h)\right) \mathrm{d}x\right\} h \ \mathrm{d}h\\
&\overset{(d)}{\ge} \int_0^\infty \mathcal{V}\left\{\liminf_{\eta\to\infty}\int_{-\infty}^\infty P_e\left[\eta; x, x+h, p_0(x,h), p_1(x,h)\right]\right.\\
&\qquad \left. \cdot \left(f_X(x)+f_X(x+h)\right) \mathrm{d}x\right\} h \ \mathrm{d}h\\
&\overset{(e)}{\ge} \int_0^\infty \mathcal{V}\Big\{\int_{-\infty}^\infty \min\left\{f_X(x), f_X(x+h)\right\} \mathrm{d}x\Big\} \ h \ \mathrm{d}h, \quad (20)
\end{aligned}$$

where: (c) follows by using Fatou's lemma; (d) follows from (17); and (e) follows by using Fatou's lemma, (16) and assumption **A2b**.

Combining the upper bound on the limit in (19) and the lower bound on the limit in (20) we arrive at

$$\begin{aligned}
&\lim_{\eta\to\infty} 2\overline{\mathsf{LB}_{\mathsf{ZZ}}}(X,\eta)\\
&= \int_0^\infty \mathcal{V}\Big\{\int_{-\infty}^\infty \min\left\{f_X(x), f_X(x+h)\right\} \mathrm{d}x\Big\} \ h \ \mathrm{d}h\\
&= \frac{1}{2}\int_{-\infty}^\infty \mathcal{V}\Big\{\int_{-\infty}^\infty \min\left\{f_X(x), f_X(x+h)\right\} \mathrm{d}x\Big\} \ |h| \ \mathrm{d}h,
\end{aligned}$$

where in the last step we have used that

$$\begin{aligned}
&\int_{-\infty}^\infty \min\left\{f_X(x), f_X(x+h)\right\} \ \mathrm{d}x\\
&\quad = \int_{-\infty}^\infty \min\left\{f_X(x), f_X(x-h)\right\} \ \mathrm{d}x.
\end{aligned}$$

This concludes the proof of the limit for $\overline{\mathsf{LB}_{\mathsf{ZZ}}}(X,\eta)$. To obtain the limit for $\mathsf{LB}_{\mathsf{ZZ}}(X,\eta)$, we simply drop the valley-filling function. With this we arrive at

$$\begin{aligned}
&\lim_{\eta\to\infty} 2\mathsf{LB}_{\mathsf{ZZ}}(X,\eta)\\
&= \frac{1}{2}\int_{-\infty}^\infty \int_{-\infty}^\infty \min\left\{f_X(x), f_X(x+h)\right\} |h| \ \mathrm{d}x \ \mathrm{d}h\\
&= \frac{1}{2}\int_{-\infty}^\infty \int_{-\infty}^\infty \min\left\{f_X(x), f_X(y)\right\} |y-x| \ \mathrm{d}x \ \mathrm{d}y.
\end{aligned}$$

This concludes the proof of Theorem 2.

Appendix B
Examples for the High-Noise Regime

A. Example 1

We will show that $\mathbb{V}(X) = \mathsf{Var}(X)$, which, in view of (15), will also characterize $\overline{\mathbb{V}}(X)$. By substituting $f_X(x) = g(|x|)$ inside the expression of $\mathbb{V}(X)$ in Theorem 2, we arrive at

$$\begin{aligned}
4\ \mathbb{V}(X) &= \int_{-\infty}^\infty \int_{-\infty}^\infty \min\{g(|x|), g(|y|)\} \ |y-x| \ \mathrm{d}x \ \mathrm{d}y\\
&= 4\int_0^\infty \int_{-y}^y g(|y|) \ |y-x| \ \mathrm{d}x \ \mathrm{d}y\\
&= 8\int_0^\infty g(|y|) \ y^2 \ \mathrm{d}y \overset{(a)}{=} 4 \ \mathsf{Var}(X),
\end{aligned}$$

where (a) follows since $\mathbb{E}[X] = 0$ from the structure of $f_X(x)$.

B. Example 2

By using the expression of $\mathbb{V}(X)$ in Theorem 2, we arrive at

$$\begin{aligned}
4\ \mathbb{V}(X) &= \int_a^b \int_a^b \min\{f_X(x), f_X(y)\} \ |y-x| \ \mathrm{d}x \ \mathrm{d}y\\
&= 2\int_a^b \int_a^y f_X(y) \ |y-x| \ \mathrm{d}x \ \mathrm{d}y\\
&= \int_a^b f_X(y) \ |y-a|^2 \ \mathrm{d}y\\
&= \mathbb{E}[(X-a)^2] = \mathsf{Var}(X) + (a-\mathbb{E}[X])^2.
\end{aligned}$$

By using the expression of $\overline{\mathbb{V}}(X)$ in Theorem 2, we have

$$\begin{aligned}
2\overline{\mathbb{V}}(X) &\overset{(a)}{=} \int_0^\infty \mathcal{V}\Big\{\int_a^b \min\left\{f_X(x), f_X(x+h)\right\} \mathrm{d}x\Big\} \ h \ \mathrm{d}h\\
&\overset{(b)}{=} \int_0^\infty \mathcal{V}\Big\{\int_a^{\max\{a, b-h\}} f_X(x+h) \mathrm{d}x\Big\} \ h \ \mathrm{d}h\\
&= \int_0^\infty \mathcal{V}\Big\{\int_{a+h}^{\max\{a+h, b\}} f_X(x) \mathrm{d}x\Big\} \ h \ \mathrm{d}h\\
&\overset{(c)}{=} \int_0^\infty \mathcal{V}\Big\{\mathbb{P}[X \ge a+h]\Big\} \ h \ \mathrm{d}h\\
&\overset{(d)}{=} \int_0^\infty \mathbb{P}[X \ge a+h] \ h \ \mathrm{d}h\\
&\overset{(e)}{=} \frac{1}{2}\mathbb{E}[(X-a)^2] = \frac{1}{2}\left(\mathsf{Var}(X) + (a-\mathbb{E}[X])^2\right),
\end{aligned}$$

where the labeled equalities follow from: (a) using the fact that X is supported on (a,b); (b) the assumption that $f_X(x)$ is a non-increasing function; (c) the fact that X is supported on (a,b) and hence, we can drop the upper limit; (d) the fact that $h \mapsto \mathbb{P}[X \ge a+h]$ is a non-increasing function and hence, the valley-filling function can be dropped; and (e) the following alternative representation of the second moment of non-negative random variables: for a random variable $U \ge 0$, we have that $\mathbb{E}[U^2] = 2\int_0^\infty \mathbb{P}[U \ge h] \ h \ \mathrm{d}h$.

REFERENCES

[1] K. L. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Trees, "Extended Ziv-Zakai lower bound for vector parameter estimation," *IEEE Transactions on Information Theory*, vol. 43, no. 2, pp. 624–637, 1997.

[2] K. Bell, Y. Ephraim, and H. Van Trees, "Explicit Ziv-Zakai lower bound for bearing estimation," *IEEE Transactions on Signal Processing*, vol. 44, no. 11, pp. 2810–2824, 1996.

[3] D. Guo, Y. Wu, S. Shamai (Shitz), and S. Verdú, "Estimation in Gaussian noise: Properties of the minimum mean-square error," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2371–2385, 2011.

[4] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., 1993.

[5] A. Bhatt, J.-T. Huang, Y.-H. Kim, J. J. Ryu, and P. Sen, "Variations on a theme by Liu, Cuff, and Verdú: The power of posterior sampling," in *2018 IEEE Information Theory Workshop (ITW)*, 2018, pp. 1–5.

[6] E. Weinstein and A. J. Weiss, "A general class of lower bounds in parameter estimation," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 338–342, 1988.

[7] H. Van Trees, K. Bell, and Z. Tian, *Detection, Estimation, and Modulation Theory Part I*. John Wiley & Sons, New York, 1968.

[8] B. Bobrovsky and M. Zakai, "A lower bound on the estimation error for certain diffusion processes," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 45–52, 1976.

[9] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 1084–1093, 1997.

[10] B.-Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai, "Some classes of global Cramér-Rao bounds," *The Annals of Statistics*, pp. 1421–1438, 1987.

[11] H. L. Van Trees and K. L. Bell, "Bayesian bounds for parameter estimation and nonlinear filtering/tracking," *AMC*, vol. 10, no. 12, pp. 10–1109, 2007.

[12] A. Renaux, P. Forster, P. Larzabal, C. D. Richmond, and A. Nehorai, "A fresh look at the Bayesian bounds of the Weiss-Weinstein family," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5334–5352, 2008.

[13] A. Dytso, M. Fauß, A. M. Zoubir, and H. V. Poor, "MMSE bounds for additive noise channels under Kullback–Leibler divergence constraints on the input distribution," *IEEE Transactions on Signal Processing*, vol. 67, no. 24, pp. 6352–6367, 2019.

[14] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Transactions on Information Theory*, vol. 15, no. 3, pp. 386–391, 1969.

[15] L. P. Seidman, "Performance limitations and error calculations for parameter estimation," *Proceedings of the IEEE*, vol. 58, no. 5, pp. 644–652, 1970.

[16] D. Chazan, M. Zakai, and J. Ziv, "Improved lower bounds on signal parameter estimation," *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 90–93, 1975.

[17] S. Bellini and G. Tartara, "Bounds on error in signal parameter estimation," *IEEE Transactions on Communications*, vol. 22, no. 3, pp. 340–342, 1974.

[18] E. Weinstein, "Relations between Belini-Tartara, Chazan-Zakai-Ziv, and Wax-Ziv lower bounds," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 342–343, 1988.

[19] W. Yihong and S. Verdú, "MMSE dimension," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4857–4879, 2011.

[20] H. Steendam and M. Moeneclaey, "Low-SNR limit of the Cramer-Rao bound for estimating the carrier phase and frequency of a PAM, PSK, or QAM waveform," *IEEE Communications letters*, vol. 5, no. 5, pp. 218–220, 2001.

[21] F. Rice, B. Cowley, B. Moran, and M. Rice, "Cramer-Rao lower bounds for QAM phase and frequency estimation," *IEEE Transactions on Communications*, vol. 49, no. 9, pp. 1582–1591, 2001.

[22] W. Xu, A. B. Baggeroer, and C. D. Richmond, "Bayesian bounds for matched-field parameter estimation," *IEEE Transactions on Signal Processing*, vol. 52, no. 12, pp. 3293–3305, 2004.

[23] A. Dytso and H. V. Poor, "Estimation in Poisson noise: Properties of the conditional mean estimator," *IEEE Transactions on Information Theory*, vol. 66, no. 7, pp. 4304–4323, 2020.

[24] D. Guo, S. Shamai (Shitz), and S. Verdú, "Mutual information and conditional mean estimation in Poisson channels," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1837–1849, 2008.

[25] T. Cover and J. Thomas, *Elements of Information Theory: Second Edition*. Wiley, 2006.

[26] J. Jiao, K. Venkat, and T. Weissman, "Relations between information and estimation in discrete-time Lévy channels," *IEEE Transactions on Information Theory*, vol. 63, no. 6, pp. 3579–3594, 2017.

[27] M. Fauß, A. Dytso, and H. V. Poor, "A variational interpretation of the Cramér–Rao bound," *Signal Processing*, vol. 182, p. 107917, 2021.

[28] A. Dembo, T. Cover, and J. Thomas, "Information theoretic inequalities," *IEEE Transactions on Information Theory*, vol. 37, no. 6, pp. 1501–1518, 1991.