A Spectral Representation of the Simple Hypothesis Testing and Refined Approximations

Barış Nakiboğlu
Department of Electrical and Electronics Engineering
Middle East Technical University
Ankara, Türkiye
0000-0001-7737-5423

2026 14th Iran Workshop on Communication and Information Theory (IWCIT) | 979-8-3195-0660-3/26/$31.00 ©2026 IEEE | DOI: 10.1109/IWCIT71404.2026.11768870

***Abstract*—For the simple hypothesis testing with randomized tests, the infimum type II error probability/volume β for a given Type I error probability/volume threshold ϵ is expressed as a Riemann integral of the cumulative distribution function of the log-likelihood ratio: $\beta(\epsilon)=\int_{-\infty}^{\infty}e^{-\tau}|F(\tau)-\epsilon|^{+}d\tau$ for all $\epsilon\geq 0$. Measure change identities for $\beta(\cdot)$ are derived under various monotonicity assumptions on the functions relating the Radon-Nikodym derivative of an auxiliary measure to those of the null and alternative σ-finite measures. These expressions yield state-of-the-art approximations for $\beta(\cdot)$ with briefer analyses.**

I. Introduction

The simple hypothesis testing problem is a fundamental problem motivated by diverse applications, particularly in information theory, see [1]–[5]. Its significance for information transmission was already evident by the time of Strassen's seminal work [6]. As pointed out in [6], and employed in [7], lossless source coding provides an operational justification for considering the cases when the measure associated with the alternative hypothesis is a σ-finite measure rather than a probability measure. Specifically, the problem of finding the minimum cardinality of a subset of a countably infinite set that has all but ϵ of the probability mass is a hypothesis testing problem against the counting measure $\sharp$, with deterministic tests. While the counting measure $\sharp$ is not finite for countably infinite sets, it remains σ-finite for any countable set. Beyond the lossless source coding, simple hypothesis testing is instrumental in characterizing the behavior of the error probability in the channel coding problem, not only at rates near capacity [6], but also at other rates [8]–[12].

For any pair of probability measures, Neyman-Pearson lemma determines the minimum type II error (i.e., missed detection) probability for a given type I error (i.e., false alarm) probability for the simple hypothesis testing with randomized tests. Thus Neyman-Pearson lemma implicitly determine type II error probability as a function of type I error probability and vice-versa. Some of the implications of this fact were explored in [13, Theorems 21 and 22], [14, Theorem 21], [15, Lemmas 1 and 2], [16, Theorem 21], [17] for various parameterizations and under various hypothesis.

We call the infimum type II error volume as a function of type I error volume the deterministic test Neyman-Pearson function (dNP-function) and denote it by $\beta(\cdot\|_{w}^{q})$ when the measures associated with the alternative and null hypotheses are q and w, respectively. If the randomized tests are permitted then we call it the randomized test Neyman-Pearson function (rNP-function) and denote it by $\beta(\cdot|_{w}^{q})$. That is for any pair of σ-finite measures (w,q) on the measurable space $(\mathcal{Y},\mathcal{Y})$, we define dNP-function $\beta(\cdot\|_{w}^{q}):[0,\infty)\mapsto[0,\infty]$ and rNP-function $\beta(\cdot|_{w}^{q}):[0,\infty)\mapsto[0,\infty]$ as

$$\beta(\epsilon\|_{w}^{q}):=\inf_{\mathsf{T}\in\mathcal{T}^{d}(\nu):\int\mathsf{T}\mathrm{d}w\leq\epsilon}\int(1-\mathsf{T})\mathrm{d}q\quad\forall\epsilon\in[0,\infty),\tag{1a}$$

$$\beta(\epsilon|_{w}^{q}):=\inf_{\mathsf{T}\in\mathcal{T}(\nu):\int\mathsf{T}\mathrm{d}w\leq\epsilon}\int(1-\mathsf{T})\mathrm{d}q\quad\forall\epsilon\in[0,\infty),\tag{1b}$$

where ν is any σ-finite reference measure in which both w and q are absolutely continuous (i.e., for which $w\prec\nu$ and $q\prec\nu$), $\mathcal{T}^{d}(\nu)$ is the set of all ν-measurable deterministic tests on $\mathcal{Y}$ (i.e., ν-measurable function of the form $\mathsf{T}:\mathcal{Y}\mapsto\{0,1\}$), and $\mathcal{T}(\nu)$ is the set of all ν-measurable randomized tests on $\mathcal{Y}$ (i.e., ν-measurable function of the form $\mathsf{T}:\mathcal{Y}\mapsto[0,1]$).

To recall the connection to the lossless source coding, let $\mathcal{Y}$ be a countable set and $\mathcal{Y}$ be its power set $2^{\mathcal{Y}}$. Then for any probability mass function w on $\mathcal{Y}$, the minimum cardinality of a set $\mathcal{A}\subset\mathcal{Y}$ satisfying $w(\mathcal{A})\geq 1-\epsilon$ is $\beta\left(\epsilon\middle\|_{w}^{\sharp}\right)$, where $\sharp$ is the counting measure satisfying $\sharp(y)=1$ for all $y\in\mathcal{Y}$.

The extension of definitions of $\beta(\cdot\|_{w}^{q})$ and $\beta(\cdot|_{w}^{q})$ from the customary setting in which both w and q are probability measures, to the case when both w and q are σ-finite measures results in a subtlety: for some σ-finite measures w and q, both $\beta(\epsilon\|_{w}^{q})$ and $\beta(\epsilon|_{w}^{q})$ are infinite for all finite ϵ values. We deem those cases to be void cases. We will see that non-void cases are those cases satisfying the following condition:

$$q\left(\tfrac{\mathrm{d}q}{\mathrm{d}\nu}<\tfrac{\mathrm{d}w}{\mathrm{d}\nu}\right)+w\left(\tfrac{\mathrm{d}w}{\mathrm{d}\nu}\leq\tfrac{\mathrm{d}q}{\mathrm{d}\nu}\right)<\infty.\tag{2}$$

If σ-finite measures w and q satisfies (2), then

$$\beta(\epsilon|_{w}^{q})=\int_{0}^{\infty}\left|w\left(\tfrac{\mathrm{d}q_{\sim}}{\mathrm{d}w}>\tau\right)-\epsilon\right|^{+}\mathrm{d}\tau\qquad\forall\epsilon\geq 0,\tag{3a}$$

$$=\int_{0}^{\infty}\left|w_{\sim}\left(\tfrac{\mathrm{d}w_{\sim}}{\mathrm{d}q}\leq\tfrac{1}{\tau}\right)-\epsilon\right|^{+}\mathrm{d}\tau\qquad\forall\epsilon\geq 0,\tag{3b}$$

where $q_{\sim}$ is the w-absolutely continuous component of q and $w_{\sim}$ is the q-absolutely continuous component of w. To the best of our knowledge, (3) has not been previously reported, even for probability measures w and q satisfying $w\sim q$. Theorem 1 establishes (3) as part of a more comprehensive characterization of the rNP-function $\beta(\cdot|_{w}^{q})$. The remaining

979-8-3195-0660-3/26/$31.00 ©2026 IEEE
Authorized licensed use limited to: ULAKBIM UASL - MIDDLE EAST TECHNICAL UNIVERSITY. Downloaded on November 07,2026 at 17:35:39 UTC from IEEE Xplore. Restrictions apply.

identities in Theorem 1 are also, to our knowledge, new in both their specific form and their generality. We note, however, that expressions equivalent to (11a), (11b), and (11c) have appeared in [13]–[17] for the case when w and q are probability measures, occasionally under additional hypotheses.

In §II, we present Theorem 1, which makes the implications of the Neyman-Pearson lemma explicit and extend Neyman-Pearson lemma and its consequences to the randomized decisions between σ-finite measures in all non-void cases. In §III, we establish sufficient conditions to express $\beta(\cdot|_w^q)$ in terms of $\beta(\cdot|_v^q)$ or in terms of $\beta(\cdot|_v^w)$ for an auxiliary probability measure v. In §IV, we utilize (3) with a change of variable and apply the measure-change identities of §III to derive non-asymptotic approximations for rNP-functions. These approximations hold under a Kolmogorov distance to Gaussianity hypothesis, which is satisfied in both memoryless and Markovian settings.

II. Neyman–Pearson Lemma Revisited

Definition 1. For any pair of σ-finite measures (w,q) on the same measurable space, say $(\mathcal{Y},\mathcal{Y})$, the likelihood-ratio $\ell_{w,q}(\cdot)$ and the likelihood-ratio distribution $F_\ell(\cdot|_q^w)$ are defined as

$$\ell_{w,q}:=\begin{cases}\infty & \text{if } \frac{\mathrm{d}q}{\mathrm{d}\nu}=0,\\ \frac{\mathrm{d}w_\sim}{\mathrm{d}q} & \text{if } \frac{\mathrm{d}q}{\mathrm{d}\nu}>0,\end{cases} \tag{4a}$$

$$F_\ell(\tau|_q^w):=\int \mathbb{1}_{\{\ell_{w,q}\le\tau\}}\mathrm{d}w \qquad \forall \tau\in\mathbb{R}, \tag{4b}$$

where ν is any σ-finite reference measure satisfying $w\prec\nu$ and $q\prec\nu$, and $w_\sim$ is the q-absolutely continuous component of w.

$F_\ell(\cdot|_q^w)$ does not depend on the choice of ν. Furthermore, with the understanding that $1/0$ and $1/\infty$ stand for ∞ and 0, respectively, we have

$$\ell_{q,w}=\tfrac{1}{\ell_{w,q}} \qquad (w+q)\text{-a.e.}. \tag{5}$$

We call the cases when the likelihood-ratio distribution is a function of the form $F_\ell(\cdot|_q^w):\mathbb{R}\mapsto\{0,\gamma\}$ trivial cases. The case when $F_\ell(\tau|_q^w)=0$ for all $\tau\in\mathbb{R}$ is called the singular case because it happens iff w and q are mutually singular, i.e. $w\perp q$. We call the case when the likelihood-ratio distribution is a surjective (onto) function of the form $F_\ell(\cdot|_q^w):\mathbb{R}\mapsto\{0,\gamma\}$ the non-singular-trivial case. The non-singular-trivial case happens iff $w\not\perp q$ and $\frac{\mathrm{d}w_\sim}{\mathrm{d}q}=\gamma$ holds $w_\sim$-a.e. some $\gamma\in(0,\infty)$.

The likelihood-ratio distribution $F_\ell(\cdot|_q^w)$ might not have an inverse in the usual sense for some (w,q) pairs. Nevertheless, its monotonicity ensures that it has semi-continuous inverses for all (w,q) pairs, [18]–[24], called likelihood-ratio quantiles.

Definition 2. For any pair of σ-finite measures (w,q) on the same measurable space, the likelihood-ratio quantiles (LRQs) $F_\ell^\wedge(\cdot|_q^w):\mathbb{R}\mapsto\overline{\mathbb{R}}$ and $F_\ell^\vee(\cdot|_q^w):\mathbb{R}\mapsto\overline{\mathbb{R}}$ are defined as

$$F_\ell^\wedge(\epsilon|_q^w):=\inf\{\gamma\in\mathbb{R}:F_\ell(\gamma|_q^w)\ge\epsilon\} \qquad \forall\epsilon\in\mathbb{R}, \tag{6a}$$

$$F_\ell^\vee(\epsilon|_q^w):=\inf\{\gamma\in\mathbb{R}:F_\ell(\gamma|_q^w)>\epsilon\} \qquad \forall\epsilon\in\mathbb{R}. \tag{6b}$$

For any pair of σ-finite measures (w,q) on the same measurable space, the primitive entropy spectrum of w with respect to q, i.e., $h(\cdot|_q^w):[0,\infty)\to[0,\infty]$, is defined as

$$h(\gamma|_q^w):=\int\Big[\tfrac{\mathrm{d}w}{\mathrm{d}\nu}\wedge\Big(\gamma\tfrac{\mathrm{d}q}{\mathrm{d}\nu}\Big)\Big]\mathrm{d}\nu \qquad \forall\gamma\in[0,\infty), \tag{7}$$

where ν is any σ-finite reference measure satisfying $w\prec\nu$ and $q\prec\nu$, [25, §31]. Furthermore, $h(\cdot|_q^w)$ is a non-decreasing function satisfying $h(\gamma|_q^w)=\gamma h(1/\gamma|_w^q)$ for all $\gamma\in(0,\infty)$. Thus

$$(1\vee\gamma)h(1|_q^w)\ge h(\gamma|_q^w)\ge(1\wedge\gamma)h(1|_q^w) \quad \forall\gamma\in(0,\infty). \tag{8}$$

Thus for any pair of σ-finite measures (w,q) and $\forall\gamma\in(0,\infty)$,

$$\text{“}h(1|_q^w)\in\mathcal{A}_\jmath \quad \text{iff} \quad h(\gamma|_q^w)\in\mathcal{A}_\jmath\text{”} \qquad \forall\jmath\in\{1,2,3\}, \tag{9}$$

where $\mathcal{A}_1=\{0\}$, $\mathcal{A}_2=(0,\infty)$, and $\mathcal{A}_3=\{\infty\}$.

In the singular case, i.e. when $w\perp q$ holds, $h(1|_q^w)=0$ and $\beta(\epsilon\|_w^q)=\beta(\epsilon|_w^q)=0$ for all $\epsilon\ge0$. More generally in the trivial case $h(\cdot|_q^w)$ has the following simple expression $h(\gamma|_q^w)=\|w_\sim\|_1\wedge(\gamma\|q_\sim\|_1)$ for all $\gamma\in[0,\infty)$. Note that a non-singular-trivial case can simultaneously be a void case. In the non-void-non-singular-trivial case, rNP-function satisfies $\beta(\epsilon|_w^q)=\|q_\sim\|_1\Big|1-\frac{\epsilon}{\|w_\sim\|_1}\Big|^+$ for all $\epsilon\ge0$.

Note that $h(1|_q^w)\le\epsilon+\beta(\epsilon|_w^q)$ for all $\epsilon\ge0$ by the definitions of $\beta(\cdot|_w^q)$ and $h(\cdot|_q^w)$ in (1) and (7). Thus we are in the void case whenever $h(1|_q^w)=\infty$, i.e., the condition in (2) is violated. On the other hand, if (2) holds, i.e., $h(1|_q^w)<\infty$, then $\beta(\cdot|_w^q)$ is finite on $(0,\infty)$, see Theorem 1 in the following.

For any pair of σ-finite measures (w,q) satisfying (2), $\epsilon\in(0,\|w_\sim\|_1)$, and $\gamma\in\Big[1/F_\ell^\vee(\epsilon|_q^w),1/F_\ell^\wedge(\epsilon|_q^w)\Big]$, let $\mathsf{T}_{\epsilon,\gamma}^{q,w}$ be

$$\mathsf{T}_{\epsilon,\gamma}^{q,w}:=\begin{cases}1 & \text{if } \ell_{q,w}>\gamma,\\ \frac{\epsilon-w(\ell_{q,w}>\gamma)}{w(\ell_{q,w}=\gamma)} & \text{if } \ell_{q,w}=\gamma \text{ and } w(\ell_{q,w}=\gamma)>0,\\ 0 & \text{otherwise.}\end{cases} \tag{10}$$

Theorem 1. *If a pair of σ-finite measures (w,q) satisfies (2), i.e., $h(1|_q^w)<\infty$, then $\beta(\cdot|_w^q)$ is a non-increasing, convex, and continuous function satisfying*

$$\beta(\epsilon|_w^q)=\sup_{\gamma\ge0}h(\gamma|_w^q)-\gamma\epsilon \qquad \forall\epsilon\ge0, \tag{11a}$$

$$=h(\gamma_\epsilon|_w^q)-\gamma_\epsilon\epsilon \qquad \forall\epsilon>0, \tag{11b}$$

$$=\int_0^{\gamma_\epsilon}\Big[w\Big(\tfrac{\mathrm{d}q_\sim}{\mathrm{d}w}>\tau\Big)-\epsilon\Big]\mathrm{d}\tau \qquad \forall\epsilon\ge0, \tag{11c}$$

$$=\int_0^{\infty}\Big|w\Big(\tfrac{\mathrm{d}q_\sim}{\mathrm{d}w}>\tau\Big)-\epsilon\Big|^+\mathrm{d}\tau \qquad \forall\epsilon\ge0, \tag{11d}$$

$$\mathrm{D}_-\beta(\epsilon|_w^q)=\tfrac{-1}{F_\ell^\wedge(\epsilon|_q^w)} \qquad \forall\epsilon>0, \tag{11e}$$

$$\mathrm{D}_+\beta(\epsilon|_w^q)=\tfrac{-1}{F_\ell^\vee(\epsilon|_q^w)} \qquad \forall\epsilon\ge0, \tag{11f}$$

for any $\gamma_\epsilon\in\Big[\frac{1}{F_\ell^\vee(\epsilon|_q^w)},\frac{1}{F_\ell^\wedge(\epsilon|_q^w)\vee0}\Big]$. Furthermore, the restriction of $\beta(\cdot|_w^q)$ to $(0,\|w_\sim\|_1)$ is an invertible function with the image $(0,\|q_\sim\|_1)$ and there exists a $\mathsf{T}_\epsilon\in\mathcal{T}(\nu)$ satisfying both $\int\mathsf{T}_\epsilon\mathrm{d}w=\epsilon$ and $\int(1-\mathsf{T}_\epsilon)\mathrm{d}q=\beta(\epsilon|_w^q)$ for all $\epsilon\in[0,\|w_\sim\|_1]$.

$$\beta(\epsilon|_w^q)=\int_\epsilon^{\|w_\sim\|_1}g(\tau)\mathrm{d}\tau \qquad \forall\epsilon\in[0,\|w_\sim\|_1], \tag{11g}$$

$$\partial\beta(\epsilon|_w^q)\subset(-\infty,0) \qquad \forall\epsilon\in(0,\|w_\sim\|_1), \tag{11h}$$

for any $g:(0,\|w_\sim\|_1)\to(0,\infty)$ s.t. $\frac{1}{F_\ell^\vee(\cdot|_q^w)}\le g(\cdot)\le\frac{1}{F_\ell^\wedge(\cdot|_q^w)}$. If $q_\sim$ is a finite measure, i.e., $\|q_\sim\|_1<\infty$, then

$$\beta(\epsilon|_w^q)=\|q_\sim\|_1-\int_0^\epsilon g(\tau)\mathrm{d}\tau \quad \forall\epsilon\ge0. \tag{11j}$$

Authorized licensed use limited to: ULAKBIM UASL - MIDDLE EAST TECHNICAL UNIVERSITY. Downloaded on November 07,2026 at 17:35:39 UTC from IEEE Xplore. Restrictions apply.

Note that $h(\gamma|_w^q) = q(\ell_{w,q} > 1/\gamma) + \gamma w(\ell_{w,q} \le 1/\gamma)$ for all γ. Hence the auxiliary variable γ in (11a) is the likelihood ratio threshold in the classical Neyman-Pearson lemma. One can prove Theorem 1 using this observation, together with the elementary properties of convex functions and a generalization of the tail expectation formula for expected value of a non-negative variable. A complete proof of Theorem 1 is presented in [26, §3].

In (11c) and (11d), the '$>$' signs can be replaced by '$\ge$' signs because monotonic functions have at most countably many discontinuities by [29, Theorem 4.30] and any countable set has zero Lebesgue measure. Hence (4) and (5) imply

$$\beta(\epsilon|_w^q) = \int_0^{\gamma_\epsilon} \left[F_\ell\left(\tfrac{1}{\tau}\middle|_q^w\right) - \epsilon\right] \mathrm{d}\tau \qquad \forall \epsilon \ge 0, \tag{12a}$$

$$= \int_0^{\infty} \left|F_\ell\left(\tfrac{1}{\tau}\middle|_q^w\right) - \epsilon\right|^+ \mathrm{d}\tau \qquad \forall \epsilon \ge 0. \tag{12b}$$

As an immediate consequence of Theorem 1, we can express the convex conjugate $\beta^*(\cdot|_w^q)$ of $\beta(\cdot|_w^q)$ in terms of the primitive entropy spectrum, as follows

$$\beta^*(\lambda|_w^q) := \sup_{x \ge 0} \lambda x - \beta(x|_w^q) = \begin{cases} -h(-\lambda|_w^q) & \text{if } \lambda \le 0, \\ \infty & \text{if } \lambda > 0. \end{cases} \tag{13}$$

Furthermore, rNP-function $\beta(\cdot|_w^q) : (0, \|w_\sim\|_1) \mapsto (0, \|q_\sim\|_1)$ and rNP-function $\beta\left(\cdot\middle|_q^w\right) : (0, \|q_\sim\|_1) \mapsto (0, \|w_\sim\|_1)$ are inverses of one another. That is

$$\beta\left(\beta(\epsilon|_w^q)\middle|_q^w\right) = \epsilon \qquad \forall \epsilon \in (0, \|w_\sim\|_1), \tag{14a}$$

$$\beta\left(\beta\left(\varepsilon\middle|_q^w\right)\middle|_w^q\right) = \varepsilon \qquad \forall \varepsilon \in (0, \|q_\sim\|_1). \tag{14b}$$

Remark 1. $\beta(\cdot|_w^q) \le \beta(\cdot\|_w^q)$ because $\mathcal{T}^d(\nu) \subset \mathcal{T}(\nu)$. When q is the counting measure $\sharp$, we have an exact expression for $\beta(\cdot\|_w^q)$: According to (1a), to calculate $\beta\left(\epsilon\middle\|_w^\sharp\right)$ we need to find an $\mathcal{A} \subset \mathcal{Y}$ with fewest number of elements satisfying $w(\mathcal{Y} \setminus \mathcal{A}) \le \epsilon$; the cardinality of this set is $\beta\left(\epsilon\middle\|_w^\sharp\right)$. For ϵ values for which $\mathcal{A}$ satisfies $w(\mathcal{Y} \setminus \mathcal{A}) = \epsilon$, we have $\beta\left(\epsilon\middle\|_w^\sharp\right) = \beta\left(\epsilon\middle|_w^\sharp\right)$. On the other hand, $\beta\left(\epsilon\middle|_w^\sharp\right) < \beta\left(\epsilon\middle\|_w^\sharp\right) \le \beta\left(\epsilon\middle|_w^\sharp\right) + 1$ whenever $w(\mathcal{Y} \setminus \mathcal{A}) < \epsilon$. Since $\beta\left(\cdot\middle\|_w^\sharp\right)$ is integer valued, we get

$$\beta\left(\epsilon\middle\|_w^\sharp\right) = \lceil \beta\left(\epsilon\middle|_w^\sharp\right) \rceil \qquad \forall \epsilon \ge 0. \tag{15}$$

III. Measure Change Identities

Assumption 1. w and q are σ-finite measures satisfying (2) and $w \not\perp q$, i.e., satisfying $h\left(1\middle|_q^w\right) \in (0, \infty)$, and v is a probability measure satisfying $v \sim w_\sim$ and

$$\tfrac{\mathrm{d}v}{\mathrm{d}q} = f\left(\tfrac{\mathrm{d}w_\sim}{\mathrm{d}q}\right) \qquad w_\sim\text{-a.e.}, \tag{16a}$$

for a function f of the form $f : (0, \infty) \mapsto (0, \infty)$, where $w_\sim$ is the q-absolutely continuous component of w. Furthermore, $\pi(\epsilon, v)$ is defined in terms of $\mathsf{T}_{\epsilon,\gamma}^{q,w}$ defined in (10) as

$$\pi(\epsilon, v) := \begin{cases} 0 & \text{if } \epsilon = 0, \\ \displaystyle\int \mathsf{T}_{\epsilon,\gamma}^{q,w} \mathrm{d}v & \text{if } \epsilon \in (0, \|w_\sim\|_1), \\ 1 & \text{if } \epsilon \ge \|w_\sim\|_1, \end{cases} \tag{16b}$$

for any $\gamma \in \left[1/F_\ell^\vee\left(\epsilon\middle|_q^w\right), 1/F_\ell^\wedge\left(\epsilon\middle|_q^w\right)\right]$.

The Csiszar-dual $\widetilde{f} : (0, \infty) \mapsto \mathbb{R}$ of a function f of the form $f : (0, \infty) \mapsto \mathbb{R}$ is defined as, see [13, (2)], [30, (29)].

$$\widetilde{f}(x) := x\, f(1/x) \qquad \forall x \in (0, \infty). \tag{17}$$

Theorem 2. *Let measures w, q, v and function $f(\cdot)$ satisfy Assumption 1. The Csiszar-dual $\widetilde{f}$ of f defined in* (17) *satisfies*

$$\tfrac{\mathrm{d}v}{\mathrm{d}w} = \widetilde{f}\left(\tfrac{\mathrm{d}q_\sim}{\mathrm{d}w}\right) \qquad q_\sim\text{-a.e.} \tag{18a}$$

If $f : (0, \infty) \mapsto (0, \infty)$ is increasing and bijective, then

$$\beta(\epsilon|_w^q) = \beta(\pi(\epsilon, v)|_v^q) \qquad \forall \epsilon \ge 0, \tag{18b}$$

If $f : (0, \infty) \mapsto (0, \infty)$ is decreasing and bijective, and $\|q_\sim\|_1$ is finite, then

$$\beta(\epsilon|_w^q) = \|q_\sim\|_1 - \beta(1 - \pi(\epsilon, v)|_v^q) \quad \forall \epsilon \ge 0. \tag{18c}$$

If $\widetilde{f} : (0, \infty) \mapsto (0, \infty)$ is increasing and bijective, then

$$\epsilon = \beta(1 - \pi(\epsilon, v)|_v^w) \qquad \forall \epsilon \in [0, \|w_\sim\|_1]. \tag{18d}$$

If $\widetilde{f} : (0, \infty) \mapsto (0, \infty)$ is decreasing and bijective, and $\|w_\sim\|_1$ is finite, then

$$\epsilon = \|w_\sim\|_1 - \beta(\pi(\epsilon, v)|_v^w) \qquad \forall \epsilon \in [0, \|w_\sim\|_1]. \tag{18e}$$

Proof of Theorem 2. First note that,

$$\tfrac{\mathrm{d}v}{\mathrm{d}w} = \tfrac{\mathrm{d}q_\sim}{\mathrm{d}w}\tfrac{\mathrm{d}v}{\mathrm{d}q_\sim} \quad q_\sim\text{-a.e.} \quad \text{and} \quad \tfrac{1}{\frac{\mathrm{d}q_\sim}{\mathrm{d}w}} = \tfrac{\mathrm{d}w_\sim}{\mathrm{d}q} \quad q_\sim\text{-a.e.}$$

Then (18a) follows from (16a) and (17). On the other hand, [26, Lemma 5] and Theorem 1 imply that $\exists \gamma_\epsilon \in (0, \infty)$ s.t.

$$\int (1 - \mathsf{T}_{\epsilon,\gamma_\epsilon}^{q,w}) \mathrm{d}q = \beta(\epsilon|_w^q) \qquad \forall \epsilon \in (0, \|w_\sim\|_1), \tag{19}$$

$$\int \mathsf{T}_{\epsilon,\gamma_\epsilon}^{q,w} \mathrm{d}w = \epsilon \qquad \forall \epsilon \in (0, \|w_\sim\|_1). \tag{20}$$

Let us first consider the case when $f(\cdot)$ is an increasing and bijective function.

$$\begin{aligned} h\left(\tfrac{1}{f(1/\gamma_\epsilon)}\middle|_v^q\right) &\overset{(a)}{=} q\left(\tfrac{\mathrm{d}v}{\mathrm{d}q} > f(1/\gamma_\epsilon)\right) + \tfrac{1}{f(1/\gamma_\epsilon)} v\left(\tfrac{\mathrm{d}v}{\mathrm{d}q} \le f(1/\gamma_\epsilon)\right) \\ &\overset{(b)}{=} q\left(\tfrac{\mathrm{d}w_\sim}{\mathrm{d}q} > \tfrac{1}{\gamma_\epsilon}\right) + \tfrac{1}{f(1/\gamma_\epsilon)} v\left(\tfrac{\mathrm{d}w_\sim}{\mathrm{d}q} \le \tfrac{1}{\gamma_\epsilon}\right) \\ &\overset{(c)}{=} \beta(\epsilon|_w^q) + \tfrac{1}{f(1/\gamma_\epsilon)} \pi(\epsilon, v), \end{aligned}$$

where (a) follows from (7), (b) follows from (16a) because $f(\cdot)$ is increasing and invertible, (c) follows from (10), (16b), and (19). Thus $\beta(\pi(\epsilon, v)|_v^q) \ge \beta(\epsilon|_w^q)$ by (11a). On the other hand $\beta(\pi(\epsilon, v)|_v^q) \le \beta(\epsilon|_w^q)$ by (1b), (16b), and (19) because $\mathsf{T}_{\epsilon,\gamma_\epsilon}^{q,w} \in \mathcal{T}(\nu)$ by [26, Lemma 5]. Thus (18b) holds for all $\epsilon \in (0, \|w_\sim\|_1)$. The $\epsilon = 0$ and $\epsilon \ge \|w_\sim\|_1$ cases, are evident as a result of $v \sim w_\sim$ and Theorem 1. Proofs of the remaining cases are similar, see [28]. □

Theorem 2 is inspired by the information theorists approach to large deviations analysis, which can be traced back to [8]–[12]. The tilted probability measures defined in the following is considered in [8, Theorem 5] for $\alpha \in (0, 1)$. For all $\alpha \in \mathbb{R} \setminus \{0, 1\}$, both $f(\cdot)$ and $\widetilde{f}(\cdot)$ associated with v_α satisfy the monotonicity assumptions of Theorem 2.

Authorized licensed use limited to: ULAKBIM UASL - MIDDLE EAST TECHNICAL UNIVERSITY. Downloaded on November 07,2026 at 17:35:39 UTC from IEEE Xplore. Restrictions apply.

Definition 3. For any pair of σ-finite measures (w, q) satisfying $w \not\perp q$, let $\mu\big(\cdot\big|_q^w\big) : \mathbb{R} \mapsto (-\infty, \infty]$ be

$$\mu\big(\alpha\big|_q^w\big) := \ln \int \big(\tfrac{\mathrm{d}w_\sim}{\mathrm{d}q_\sim}\big)^\alpha \mathrm{d}q_\sim \qquad \forall \alpha \in \mathbb{R}, \tag{21a}$$

$$\mathcal{I}_q^w := \{\alpha \in \mathbb{R} : \mu\big(\alpha\big|_q^w\big) < \infty\}. \tag{21b}$$

The tilted probability measure v_α is defined for all $\alpha \in \mathcal{I}_q^w$, as a $w_\sim$-equivalent measure via the following relation

$$\tfrac{\mathrm{d}v_\alpha}{\mathrm{d}q} := e^{-\mu\left(\alpha\big|_q^w\right)} \big(\tfrac{\mathrm{d}w_\sim}{\mathrm{d}q}\big)^\alpha \qquad w_\sim\text{-a.e.}, \tag{21c}$$

Note that $\mu\big(\cdot\big|_q^w\big)$ is a convex function as a result of Hölder's inequality [31, Theorem 2.11.1]. Thus $\mathcal{I}_q^w$ is an interval. Furthermore, the tilted probability measure v_α also satisfies

$$\tfrac{\mathrm{d}v_\alpha}{\mathrm{d}w} = e^{-\mu\left(\alpha\big|_q^w\right)} \big(\tfrac{\mathrm{d}q_\sim}{\mathrm{d}w}\big)^{1-\alpha} \qquad q_\sim\text{-a.e..} \tag{22}$$

Remark 2. In the non-singular-trivial case $\exists \gamma \in (0, \infty)$ such that $\mu\big(\alpha\big|_q^w\big) = \ln\|q_\sim\|_1 + \alpha \ln \gamma$ holds $\forall \alpha \in \mathbb{R}$. Furthermore, non-singular-trivial case is non-void iff $\|q_\sim\|_1$, and hence $\|w_\sim\|_1$ is finite. Thus $\gamma = {}^{\|w_\sim\|_1}\!/_{\|q_\sim\|_1}$, $v_\alpha = {}^{w_\sim}\!/_{\|w_\sim\|_1}$, and $\pi(\epsilon, v_\alpha) = {}^{\epsilon}\!/_{\|w_\sim\|_1}$ for all $\alpha \in \mathbb{R}$ and for all $\epsilon \in [0, \|w_\sim\|_1]$ in the non-void-non-singular-trivial case.

Note that $v_\alpha \sim q_\sim$ holds and v_α satisfies Assumption 1 by definition. Furthermore,

(a) If $\alpha < 0$, then $f(\cdot)$ is bijective and decreasing.
(b) If $\alpha > 0$, then $f(\cdot)$ is bijective and increasing.
(c) If $\alpha < 1$, then $\widetilde{f}(\cdot)$ is bijective and increasing.
(d) If $\alpha > 1$, then $\widetilde{f}(\cdot)$ is bijective and decreasing.

Then for σ-finite measure w and q satisfying (2) and $w \not\perp q$, Theorem 2 implies for all $\epsilon \in [0, \|w_\sim\|_1]$, that

$$\beta(\epsilon|_w^q) = \begin{cases} \|q_\sim\|_1 - \beta\big(1 - \pi(\epsilon, v_\alpha)\big|_{v_\alpha}^q\big) & \text{if } \alpha < 0, \\ \|q_\sim\|_1 (1 - \pi(\epsilon, v_\alpha)) & \text{if } \alpha = 0, \\ \beta\big(\pi(\epsilon, v_\alpha)\big|_{v_\alpha}^q\big) & \text{if } \alpha > 0, \end{cases} \tag{23a}$$

$$\epsilon = \begin{cases} \beta\big(1 - \pi(\epsilon, v_\alpha)\big|_{v_\alpha}^w\big) & \text{if } \alpha < 1, \\ \|w_\sim\|_1 \pi(\epsilon, v_\alpha) & \text{if } \alpha = 1, \\ \|w_\sim\|_1 - \beta\big(\pi(\epsilon, v_\alpha)\big|_{v_\alpha}^w\big) & \text{if } \alpha > 1. \end{cases} \tag{23b}$$

Furthermore, invoking (14), we can express the function $\pi(\cdot, v_\alpha) : [0, \|w_\sim\|_1] \mapsto [0, 1]$ in terms of ϵ and $\beta(\epsilon|_w^q)$.

(23) implies certain non-trivial conditions for finiteness of $\mu\big(\alpha\big|_q^w\big)$. For example for the non-void case, i.e., when (2) holds, if $\|q_\sim\|_1 = \infty$, then $\mu\big(\alpha\big|_q^w\big) = \infty$ for all $\alpha \le 0$ because otherwise (23a) would be violated. Similarly, for the non-void case, i.e., when (2) holds, if $\|w_\sim\|_1 = \infty$, then $\mu\big(\alpha\big|_q^w\big) = \infty$ for all $\alpha \ge 1$ because otherwise (23b) would be violated.

In the following we will invoke the measure change identities in (23) with the help of following observation

$$D(v_\alpha \| q) = \mu\big(\alpha\big|_q^w\big) + \alpha \mathbf{E}_{v_\alpha}[\ln \ell_{w,q}] \qquad \forall \alpha \in \mathcal{I}_q^w, \tag{24a}$$

$$D(v_\alpha \| w) = \mu\big(\alpha\big|_q^w\big) + (\alpha - 1) \mathbf{E}_{v_\alpha}[\ln \ell_{w,q}] \qquad \forall \alpha \in \mathcal{I}_q^w, \tag{24b}$$

where $\mathbf{E}_{v_\alpha}[\cdot]$ is the expectation for random variables on the probability space $(\mathcal{Y}, \mathcal{Y}, v_\alpha)$. The identities in (24) are present in [8, Theorem 5], for the case when $\mathcal{Y}$ is finite set and w and q are probability measures.

IV. Approximations for rNP-function

Both classical limit theorems and finite-sample bounds for sums of random variables can be used together with (3) to bound $\beta(\cdot|_w^q)$. To that end, let us first express (3), i.e., (12), in terms of the distribution function of the log-likelihood-ratio.

$$\mathsf{L}_{w,q} := \ln \ell_{w,q}, \tag{25a}$$

$$F_{\mathsf{L}}\big(\tau\big|_q^w\big) := \int \mathbb{1}_{\{\mathsf{L}_{w,q} \le \tau\}} \mathrm{d}w \qquad \forall \tau \in \mathbb{R}. \tag{25b}$$

Then for any pair of σ-finite measure w and q we have

$$F_{\mathsf{L}}\big(\tau\big|_q^w\big) = F_\ell\big(\exp(\tau)\big|_q^w\big) \qquad \forall \tau \in \mathbb{R}. \tag{26}$$

If w and q also satisfy (2), applying the change of variable $\tau = e^{-s}$ in the integral in (12b) and invoking (26), we get

$$\beta(\epsilon|_w^q) = \int_{-\infty}^{\infty} e^{-s} \big|F_{\mathsf{L}}\big(s\big|_q^w\big) - \epsilon\big|^+ \mathrm{d}s \qquad \forall \epsilon \ge 0. \tag{27}$$

In the following, we will consider the case when Kolmogorov distance between $F_{\mathsf{L}}\big(\cdot\big|_q^w\big)$ and a Gaussian distribution is bounded non-trivially. Let us first recall the definitions of the Gaussian probability density function $\varphi(\cdot)$, the Gaussian Q-function $Q(\cdot)$, and the Gaussian Mills ratio $\Psi(\cdot)$:

$$\varphi(\tau) := \tfrac{1}{\sqrt{2\pi}} e^{-\tau^2/2} \qquad \forall \tau \in \mathbb{R},$$

$$Q(\tau) := \int_\tau^\infty \varphi(x) \mathrm{d}x \qquad \forall \tau \in \mathbb{R},$$

$$\Psi(x) := \tfrac{Q(x)}{\varphi(x)} \qquad \forall \tau \in \mathbb{R}.$$

Theorem 3. *If w is a probability measure and q is a σ-finite measure satisfying*

$$\big|F_{\mathsf{L}}\big(\gamma - \sigma x\big|_q^w\big) - Q(x)\big| \le \delta \qquad \forall x \in \mathbb{R}, \tag{28a}$$

for a $\gamma \in \mathbb{R}$, a $\sigma > 0$, and a $\delta \in (0, 1)$, then

$$\beta(\epsilon|_w^q) \le e^{-\gamma} \tfrac{Q\left(\sigma - Q^{-1}(\epsilon - \delta)\right)}{\varphi(\sigma)\sqrt{2\pi}} \qquad \forall \epsilon \in (\delta, 1), \tag{28b}$$

$$\beta(\epsilon|_w^q) \ge e^{-\gamma} \tfrac{Q\left(\sigma - Q^{-1}(\epsilon + \delta)\right)}{\varphi(\sigma)\sqrt{2\pi}} \qquad \forall \epsilon \in (0, 1 - \delta), \tag{28c}$$

$$\beta(\epsilon|_w^q) = \tfrac{\Psi(\sigma - \tau)}{\sqrt{2\pi}} e^{-\gamma + \sigma\tau - \frac{\tau^2}{2} + \Delta} \qquad \forall \epsilon \in (\delta, 1 - \delta - Q(\sigma)), \tag{28d}$$

where $\tau = Q^{-1}(\epsilon)$ and Δ satisfies $|\Delta| \le \tfrac{\delta(\sigma + 0.5)\sqrt{2\pi}}{(1 - \delta - \epsilon) \wedge (\epsilon - \delta)}$.

Proof of Theorem 3.

$$\begin{aligned} \beta(\epsilon|_w^q) &\overset{(a)}{=} \sigma \int_{-\infty}^{\infty} e^{-(\gamma - \sigma x)} \big|F_{\mathsf{L}}\big(\gamma - \sigma x\big|_q^w\big) - \epsilon\big|^+ \mathrm{d}x, \\ &\overset{(b)}{\le} e^{-\gamma} \sigma \int_{-\infty}^{\infty} e^{\sigma x} |Q(x) + \delta - \epsilon|^+ \mathrm{d}x, \\ &\overset{(c)}{=} e^{-\gamma} \sigma \int_{-\infty}^{Q^{-1}(\epsilon - \delta)} e^{\sigma x} |Q(x) + \delta - \epsilon|^+ \mathrm{d}x, \end{aligned}$$

where (a) follows from (27) for $s = \gamma - \sigma x$, (b) follows from the hypothesis (28a), (c) follows from the fact that $Q(\cdot)$ is an invertable and decreasing function. Then (28b) follows from the identity [26, (75)] for $\lambda = Q^{-1}(\epsilon - \delta)$. (28c) is established using the lower bound on $F_{\mathsf{L}}\big(\gamma - \sigma x\big|_q^w\big)$ in (28a) by following a similar analysis. Derivation of (28d) from (28b) and (28c), is identical to the derivation [26, (73)] from [26, (72a)] and [26, (72b)], see [27] or [28] for a complete proof. □

Authorized licensed use limited to: ULAKBIM UASL - MIDDLE EAST TECHNICAL UNIVERSITY. Downloaded on November 07,2026 at 17:35:39 UTC from IEEE Xplore. Restrictions apply.

Proposition 1. *Let w_t be a probability measure, and q_t be a σ-finite measure satisfying $|D(w_t\|q_t)|<\infty$ for all $t\in\{1,\dots,n\}$ for an $n\in\mathbb{Z}_+$, w be $w=\otimes_{t=1}^n w_t$, q be $q=\otimes_{t=1}^n q_t$, and T be*

$$T=\sum\nolimits_{t=1}^{n}\mathbf{E}_{w_t}\left[\left|\ln\tfrac{\mathrm{d}w_t}{\mathrm{d}q_t}-D(w_t\|q_t)\right|^3\right].$$

If $\mathbf{V}_w[\mathsf{L}_{w,q}]>0$, i.e., the variance of $\mathsf{L}_{w,q}$ under w is positive, then (28b), (28c), (28d) *hold for $\gamma=D(w\|q)$ and $\sigma=\sqrt{\mathbf{V}_w[\mathsf{L}_{w,q}]}$, and $\delta=0.56^T/\sigma^3$ whenever $\delta\in[0,1)$.*

Proof of Proposition 1. (28a) holds as a result of Berry-Esseen theorem [32]–[34] because $\mathsf{L}_{w,q}=\sum_{t=1}^n\ln\frac{\mathrm{d}w_t}{\mathrm{d}q_t}$. We calculate the first moment and variance of $\mathsf{L}_{w,q}$ using Fubini's theorem [31, Theorem 3.4.4], giving us the expression for γ and σ^2. Then proposition follows from Theorem 3. □

Proposition 2. *For some $\alpha\in\mathbb{R}_+$ and $n\in\mathbb{Z}_+$, let (w_t,q_t) be a pair of σ-finite measures satisfying $\mu\big(\alpha\big|{}^{w_t}_{q_t}\big)\in\mathbb{R}_+$ for all $t\in\{1,\dots,n\}$, w be $w=\otimes_{t=1}^n w_t$, q be $q=\otimes_{t=1}^n q_t$, v_α be the tilted probability measure as defined in* (21c)*, and T be*

$$T=\sum\nolimits_{t=1}^{n}\mathbf{E}_{v_\alpha}\left[\left|\ln\tfrac{\mathrm{d}w_{t\sim}}{\mathrm{d}q_{t\sim}}-\mathbf{E}_{v_\alpha}\left[\ln\tfrac{\mathrm{d}w_{t\sim}}{\mathrm{d}q_{t\sim}}\right]\right|^3\right]. \tag{29}$$

Assume $D(v_\alpha\|q)$ and $D(v_\alpha\|w)$ are finite, $\sigma=\sqrt{\mathbf{V}_{v_\alpha}[\mathsf{L}_{w,q}]}$ is in $(0,\infty)$, $\delta=0.56^T/\sigma^3$ satisfies $\delta<1/4$, and let $\varepsilon(\cdot)$ be

$$\varepsilon(\phi):=e^{-D(v_\alpha\|q)}\tfrac{Q\left(\alpha\sigma-Q^{-1}(\phi)\right)}{\varphi(\alpha\sigma)\sqrt{2\pi}}\qquad\forall\phi\in(0,1). \tag{30a}$$

If $\alpha\in(0,1)$ and $\phi\in(2\delta,1\text{–}2\delta)$, then there exists a Δ satisfying both $|\Delta|\le 2\delta$ and

$$\beta\big(\varepsilon(\phi)\big|{}^{w}_{q}\big)=e^{-D(v_\alpha\|w)}\tfrac{Q\left((1-\alpha)\sigma+Q^{-1}(\phi+\Delta)\right)}{\varphi((1-\alpha)\sigma)\sqrt{2\pi}}; \tag{30b}$$

if $\alpha\in(1,\infty)$ and $\phi\in(2\delta,1-2\delta)$, then there exists a Δ satisfying both $|\Delta|\le 2\delta$ and

$$\beta\big(\varepsilon(\phi)\big|{}^{w}_{q}\big)=\|w_\sim\|_1-e^{-D(v_\alpha\|w)}\tfrac{Q\left((\alpha-1)\sigma-Q^{-1}(\phi+\Delta)\right)}{\varphi((\alpha-1)\sigma)\sqrt{2\pi}}. \tag{30c}$$

Furthermore, if $\delta\le 1/8$ and $\varepsilon=\gamma e^{-D(v_\alpha\|q)}$ holds for a γ satisfying

$$\tfrac{Q(\alpha\sigma+0.31)}{\varphi(\alpha\sigma)\sqrt{2\pi}}\le\gamma\le\tfrac{Q(\alpha\sigma-0.31)}{\varphi(\alpha\sigma)\sqrt{2\pi}}, \tag{30d}$$

then $\exists\xi$ satisfying both $|\xi|\le\frac{0.54}{1\wedge\alpha}+(|1-\alpha|\sigma+2.2)8\sqrt{2\pi}\delta$ and

$$\beta\big(\varepsilon\big|{}^{w}_{q}\big)=\begin{cases}\gamma^{\frac{\alpha-1}{\alpha}}\theta(\alpha,\sigma)e^{-D(v_\alpha\|w)+\xi} & \text{if } \alpha<1,\\ \|w_\sim\|_1-\gamma^{\frac{\alpha-1}{\alpha}}\theta(\alpha,\sigma)e^{-D(v_\alpha\|w)+\xi} & \text{if } \alpha>1.\end{cases} \tag{30e}$$

where $\theta(\alpha,\sigma)=\left(\Psi(\alpha\sigma)\right)^{\frac{1-\alpha}{\alpha}}\Psi(|\alpha-1|\sigma)(\sqrt{2\pi})^{-\frac{1}{\alpha}}$.

Proof of Proposition 2. (28b) and (28c) of Proposition 1, (21c), (23a), (24a), and (30a) imply that $|\pi(\varepsilon(\phi),v_\alpha)-\phi|\le\delta$ for all $\phi\in(\delta,1-\delta)$. Then (30b) and (30c) follow from (28b) and (28c) of Proposition 1, (22), (23b), and (24b). Derivation of (30e) from (30b) and (30c) is identical to the derivation of [26, (82h)] from [26, (82b)] and [26, (82c)] in [26]. □

For the memoryless case, Propositions 1 and 2 recover state-of-the-art bounds with briefer analysis and, occasionally, with sharper constants, see [6], [14], [15], [35] for Proposition 1 and [36]–[41] for Proposition 2.

Note that Theorem 3, does not make any explicit assumption on the structure of w and q, hence it applies to all cases in which the Kolmogorov distance of $F_{\mathsf{L}}\big(\cdot\big|{}^{w}_{q}\big)$ to the Gaussianity is bounded non-trivially.[1] Let us consider the fixed-length lossless coding of Markovian sources as an example.

Let $\mathcal{Y}_1$ be a finite set, $\mathcal{Y}_1^n$ be the n-fold the cartesian product $\mathcal{Y}_1$ with itself, w_1^n be the probability mass function of a time–homogeneous, irreducible, aperiodic k^{th} order Markov chain with transition probabilities W, and $\sharp$ be the counting measure:

$$w_1^n(y_1^n)=w_1^{\mathrm{k}}\big(y_1^k\big)\prod\nolimits_{j=k+1}^{n}W\Big(y_j\big|y_{j-k}^{j-1}\Big)\quad\forall y_1^n\in\mathcal{Y}_1^n, \tag{31a}$$

$$\sharp(y_1^n)=1\qquad\forall y_1^n\in\mathcal{Y}_1^n. \tag{31b}$$

Then [35, Theorem 21] asserts that $\exists A\in(0,\infty)$ satisfying

$$\left|F_{\mathsf{L}}\Big(\overline{D}n-\overline{\sigma}\sqrt{n}x\Big|{}^{w_1^n}_{\sharp}\Big)-Q(x)\right|\le\tfrac{A}{\sqrt{n}}\quad\forall x\in\mathbb{R},n\in\mathbb{Z}_+, \tag{32a}$$

where $\overline{D}$ and $\overline{\sigma}$ are

$$\overline{D}=\lim_{n\to\infty}\tfrac{1}{n}D(w_1^n\|\sharp), \tag{32b}$$

$$\overline{\sigma}^2=\lim_{n\to\infty}\tfrac{1}{n}\sum\nolimits_{y_1^n}\left(\ln w_1^n(y_1^n)-D(w_1^n\|\sharp)\right)^2w_1^n(y_1^n). \tag{32c}$$

Proposition 3. *For some $n\in\mathbb{Z}_+$, let w_1^n and $\sharp$ be the measures described in* (31)*. If $\overline{\sigma}>0$, then* (28b), (28c), *and* (28d) *hold for $\gamma=\overline{D}n$ and $\sigma=\overline{\sigma}\sqrt{n}$, and $\delta=A/\sqrt{n}$ for $\overline{D}$, $\overline{\sigma}$, and A described in* (32)*.*

Proof of Proposition 3. (28a) follows from (32a) and hence the proposition follows from Theorem 3. □

For the fixed-length lossless coding of time–homogeneous, irreducible, and aperiodic Markov sources, the identity in (15) and Proposition 3 improve upon best upper bound previously established in[2] [35, Theorems 19] and recover best lower bound previously established in [35, Theorems 20] with better constants.

The bound in (32a), i.e., [35, Theorem 21], was established by relying on [42]. By invoking Nagaev's original result [42] directly instead of (32a), Proposition 3 can be extended to more general abstract time-homogeneous Markov measure w and q. This extension allows for the establishment of results corresponding to Proposition 2 for the Markovian case as well, see [28].

V. Discussion

We believe Theorem 1 is not an isolated result and it is possible to develop non-asymptotic spectral representations for other information transmission problems. In light of this, a non-asymptotic variant of the information spectrum method of [4] appears more attainable than previously thought. Such spectral representations are likely to yield not only sharper results but also new connections, such as the relationship between $\beta(\cdot|{}^{q}_{w})$ and $h(\cdot|{}^{q}_{w})$ established in Theorem 1.

[1] The characterization in (27) decouples the approximation of $\beta\big(\cdot\big|{}^{q}_{w}\big)$ from assumptions on the structure of w and q and facilitates the application classical limit theorems for sums of random variables to derive sharp approximations for $\beta\big(\cdot\big|{}^{q}_{w}\big)$ with the help of Watson's Lemma, see [28].

[2] [35] points out that [35, Theorems 19] can be improved but only by "extending the involved proof of" [35, Theorems 17].

Authorized licensed use limited to: ULAKBIM UASL - MIDDLE EAST TECHNICAL UNIVERSITY. Downloaded on November 07,2026 at 17:35:39 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge, UK: Cambridge University Press, 2011.

[2] R. G. Gallager, *Information theory and reliable communication*. New York, NY: John Wiley & Sons, Inc., 1968.

[3] T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd ed. New York, NY: Wiley-Interscience, 2006.

[4] T. S. Han, *Information-Spectrum Methods in Information Theory*, ser. Stochastic Modelling and Applied Probability (Book 50). Springer Berlin Heidelberg, 2003.

[5] Y. Polyanskiy and Y. Wu, *Information Theory: From Coding to Learning*. Cambridge University Press, 2025.

[6] V. Strassen, "Asymptotische abschätzungen in Shannons Informationstheorie," in *Trans. Third Prague Conf. Inf. Theory*, 1962, pp. 689–723, (https://pi.math.cornell.edu/~pmlut/strassen.pdf).

[7] I. Csiszár and G. Longo, "On the error exponent for source coding and for testing simple statistical hypotheses," *Studia Scientiarum Mathematicarum Hungarica*, vol. 6, pp. 181–191, 1971, (http://real-j.mtak.hu/id/eprint/5457).

[8] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. I," *Information and Control*, vol. 10, no. 1, pp. 65–103, 1967.

[9] E. A. Haroutunian, "Bounds for the exponent of the probability of error for a semicontinuous memoryless channel," *Problems of Information Transmission*, vol. 4, no. 4, pp. 29–39, 1968.

[10] U. Augustin, "The capacity of the product channel," *Information and Control*, vol. 13, no. 6, pp. 594–596, 1968.

[11] E. A. Haroutunian, "Lower bound for error probability in channels with feedback," *Problems of Information Transmission*, vol. 13, no. 2, pp. 107–114, 1977.

[12] U. Augustin, "Noisy channels," Habilitation Thesis, Universität Erlangen-Nürnberg, 1978, (http://bit.ly/3bsWDgG).

[13] M. D. Reid and R. C. Williamson, "Information, divergence and risk for binary experiments," *Journal of Machine Learning Research*, vol. 12, no. 22, pp. 731–817, 2011, (arXiv:0901.0356 [stat.ML]). [Online]. Available: http://jmlr.org/papers/v12/reid11a.html

[14] Y. Polyanskiy, "Channel coding: non-asymptotic fundamental limits," Ph.D. dissertation, Princeton University, 2010.

[15] N. Elkayam and M. Feder, "Variational formulas for the power of the binary hypothesis testing problem with applications," *arXiv:1601.06810v1 [cs.IT]*, 2016.

[16] S. Verdú, "Relative information spectra with applications to statistical inference," *AIMS Mathematics*, vol. 9, no. 12, pp. 35 038–35 090, 2024. [Online]. Available: https://www.aimspress.com/article/doi/10.3934/math.20241668

[17] ——, "Correction: Relative information spectra with applications to statistical inference," *AIMS Mathematics*, vol. 10, no. 4, pp. 9322–9323, 2025. [Online]. Available: https://www.aimspress.com/article/doi/10.3934/math.2025429

[18] E. P. Klement, R. Mesiar, and E. Pap, "Quasi- and pseudo-inverses of monotone functions, and the construction of t-norms," *Fuzzy Sets and Systems*, vol. 104, no. 1, pp. 3–13, 1999. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165011498002528

[19] P. Viceník, "A note to a construction of t-norms based on pseudo-inverses of monotone functions," *Fuzzy Sets and Systems*, vol. 104, no. 1, pp. 15–18, 1999. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016501149800253X

[20] C. Feng, W. H., X. Tu, and J. Kowalski, "A note on generalized inverses of distribution function and quantile transformation," *pplied Mathematics*, no. 3, pp. 2098–2100, 2012. [Online]. Available: http://dx.doi.org/10.4236/am.2012.312A289

[21] N. Falkner and G. Teschl, "On the substitution rule for lebesgue–stieltjes integrals," *Expositiones Mathematicae*, vol. 30, no. 4, pp. 412–418, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0723086912000588

[22] P. Embrechts and M. Hofert, "A note on generalized inverses," *Mathematical Methods of Operations Research*, vol. 77, no. 3, pp. 4230–432, 2013. [Online]. Available: https://doi.org/10.1007/s00186-013-0436-7

[23] A. de La Fortelle, "Generalized inverses of increasing functions and lebesgue decomposition," *Markov Processes And Related Fields*, Aug 2020. [Online]. Available: https://minesparis-psl.hal.science/hal-03113748

[24] P. Wacker, "Please, not *another* note about generalized inverses," *arXiv:2306.06989v1 [math.CA]*, 2023. [Online]. Available: https://arxiv.org/abs/2306.06989v1

[25] P. R. Halmos, *Measure Theory*, 1st ed., ser. Graduate Texts in Mathematics. New York, NY: Springer, 1974.

[26] B. Nakiboğlu, "The spectral representations of the simple hypothesis testing problem," *arXiv:2601.09564v1 [cs.IT]*, 2026.

[27] ——, "A spectral representation of the simple hypothesis testing problem," in *2026 IEEE International Symposium on Information Theory (ISIT)*, 2026.

[28] ——, "A spectral representation of the simple hypothesis testing problem," *arXiv:2601.09564v2 [cs.IT]*, 2026.

[29] W. Rudin, *Principles of Mathematical Analysis*. New York, NY: McGraw-Hill, 1976.

[30] F. Topsøe, "Entropy and equilibrium via games of complexity," *Physica A: Statistical Mechanics and its Applications*, vol. 340, no. 1, pp. 11–31, 2004, news and Expectations in Thermostatistics. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378437104003930

[31] V. I. Bogachev, *Measure Theory*. Berlin Heidelberg: Springer-Verlag, 2007.

[32] A. C. Berry, "The accuracy of the Gaussian approximation to the sum of independent variates," *Transactions of the American Mathematical Society*, vol. 49, no. 1, pp. 122–136, 1941.

[33] C. G. Esseen, "On the Liapunoff limit of error in the theory of probability," *Arkiv för Matematik, Astronomi Och Fysik*, vol. A28, pp. 1–19, 1942.

[34] I. G. Shevtsova, "An improvement of convergence rate estimates in the Lyapunov theorem," *Doklady Mathematics*, vol. 82, no. 3, pp. 862–864, 2010.

[35] I. Kontoyiannis and S. Verdú, "Optimal lossless data compression: Non-asymptotics and asymptotics," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 777–795, 2014.

[36] B. Nakiboğlu, "A Simple Derivation of the Refined SPB for the Constant Composition Codes," in *2019 IEEE International Symposium on Information Theory (ISIT)*, Paris, France, July 2019, pp. 2659–2663.

[37] ——, "A simple derivation of the refined sphere packing bound under certain symmetry hypotheses," *Turkish Journal Of Mathematics*, vol. 44, no. 3, pp. 919–948, 2020, (arXiv:1904.12780 [cs.IT]).

[38] V. Lungu and I. Kontoyiannis, "The optimal finite-sample error probability in asymmetric binary hypothesis testing," in *2024 IEEE International Symposium on Information Theory (ISIT)*, 2024, pp. 831–836, (arXiv:2404.09605v2 [cs.IT]).

[39] A. Theocharous and I. Kontoyiannis, "Lossless data compression at pragmatic rates," in *2025 IEEE International Symposium on Information Theory (ISIT)*, 2025, pp. 1–6.

[40] A. Theocharous, L. Gavalakis, and I. Kontoyiannis, "Pragmatic lossless compression: Fundamental limits and universality," *arXiv:2501.10103v4 [cs.IT]*, 2025, (arXiv:2501.10103v4 [cs.IT]).

[41] H.-C. Cheng and B. Nakiboğlu, "Refined strong converse for the constant composition codes," in *2020 IEEE International Symposium on Information Theory (ISIT)*, June 2020, pp. 2149–2154, (arXiv:2002.11414 [cs.IT]).

[42] S. V. Nagaev, "More exact statement of limit theorems for homogeneous markov chains," *Theory of Probability & Its Applications*, vol. 6, no. 1, pp. 62–81, 1961.

Authorized licensed use limited to: ULAKBIM UASL - MIDDLE EAST TECHNICAL UNIVERSITY. Downloaded on November 07,2026 at 17:35:39 UTC from IEEE Xplore. Restrictions apply.