Refinements and Generalizations of the Shannon Lower Bound via Extensions of Kraft's Inequality

Neri Merhav

The Viterbi Faculty of Electrical & Computer Engineering
Technion—Israel Institute of Technology
Haifa, Israel

ITA 2026, San Diego, CA, U.S.A, February 8–13, 2026

The Shannon Lower Bound (SLB) as a Benchmark

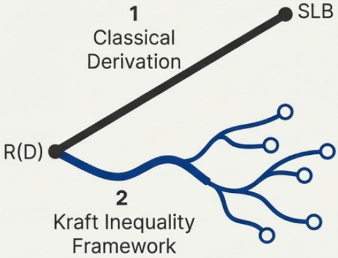
- The SLB is a cornerstome in rate-distortion theory.
- A simple lower bound. For a memoryless source U and a difference distortion measure, $d(u,v) = \rho(u-v)$:

$$R(D) \geq h(U) - \Phi(D),$$

where

$$\Phi(D) \stackrel{\triangle}{=} \sup_{\{Z:\ \mathbf{E}\{\rho(Z) \leq D\}} h(Z) \equiv \inf_{\beta \geq 0} \left\{ \beta D + \log \left[\int_{\mathcal{U}} 2^{-\beta\rho(z)} \mathsf{d}z \right] \right\}.$$

- Asymptotically tight for $D \to 0$ (Linder & Zamir '94, Koch '16).
- A useful reference for high-resolution lossy compression schemes.
- Further studies in the finite block-length regime (Kostina, '15,'16).
- Focus on the quadratic distortion metric (Gastpar & Sula '24).
- SLB via "lossy Kraft inequality" (Campbell '73, Merhav '95).



A More Revealing Path: The Kraft Inequality Approach

- The classical SLB can be derived directly through mutual information manipulations.
- However, a less-traveled path via the Kraft inequality provides a more powerful framework.
- This paper's core contribution is to demonstrate that several extended versions of the Kraft inequality pave the way to a series of non-trivial refinements and generalizations of the SLB.
- The focus is not on a simpler proof, but on the new discoveries this method enables.

The Core Methodological Tool: An Extended Kraft Integral

For a given encoder-decoder pair (ϕ_n, ψ_n) and parameters $\alpha > 1$ and $\beta \geq 0$, we define the extended Kraft integral as:

$$Z_n(\alpha, \beta) = \int_{U^n} \exp_2\{-\alpha L[\phi_n(u^n)] - \beta\rho(u^n - \psi_n(\phi_n(u^n)))\} \, \mathrm{d}u^n$$

Explanation of Terms

$L[\cdot]$: The length of the compressed codeword.
$\rho(\cdot)$: The distortion incurred.
α, β: Parameters weighting the rate and distortion costs.

Key Result (Lemma 1): For any uniquely decodable (UD) code, this integral is bounded.

$$Z_n(\alpha, \beta) \leq \left[\int_U 2^{-\beta\rho(z)} dz\right]^n$$

This inequality connects the code's performance (rate and distortion) to a fundamental property of the distortion measure.

Recovering the known SLB in Different Forms

$$\begin{aligned} & \left[\int_{\mathcal{U}} 2^{-\beta\rho(z)}\mathrm{d}z\right]^n \geq Z_n(\alpha,\beta) \\ = & \int_{\mathcal{U}^n} \exp_2\left\{-\alpha L[\phi_n(u^n)] - \beta\rho(u^n - \psi_n(\phi_n(u^n)))\right\} \\ = & \ \mathbf{E}\left\{\exp_2\left[-\alpha L[\phi_n(U^n)] - \beta\rho(U^n - \psi_n(\phi_n(U^n))) - \log P(U^n)\right]\right\} \\ \geq & \ \exp_2\left[-\alpha\mathbf{E}\{L[\phi_n(U^n)]\} - \beta\mathbf{E}\{\rho(U^n - \psi_n(\phi_n(U^n)))\} + h(U^n)\right], \end{aligned}$$

and so, rearranging terms, we get

$$\alpha\mathbf{E}\{L[\phi_n(U^n)]\} + \beta\mathbf{E}\{\rho(U^n - \psi_n(\phi_n(U^n)))\} \geq h(U^n) - n\log\left[\int_{\mathcal{U}} 2^{-\beta\rho(z)}\mathrm{d}z\right].$$

which yields

$$\mathbf{E}\{L[\phi_n(U^n)]\} + \beta\mathbf{E}\{\rho(U^n - \psi_n(\phi_n(U^n)))\} \geq h(U^n) - n\log\left[\int_{\mathcal{U}} 2^{-\beta\rho(z)}\mathrm{d}z\right].$$

Invoking the distortion constraint, moving to the r.h.s., and maximizing over β yields the classical SLB, $\mathbf{E}\{L[\phi_n(U^n)]\} \geq h(U^n) - n\Phi(D)$.

Application 1: A Sharper Bound for One-to-One Codes

Scenario: We relax the constraint from uniquely decodable (UD) codes to the broader class of one-to-one codes at the block level.

Method: A variation of the Kraft inequality (Lemma 2) is used, which introduces a denominator term reflecting the relaxed constraint.

$$Z_{n,1-1}(\alpha, \beta) \leq \frac{\left[\int_U 2^{-\beta\rho(z)} dz\right]^n}{2^{\alpha-1} - 1}.$$

Result: Following the same derivation path, this new term leads to a refined lower bound on the rate.

$$\frac{\mathbb{E}[L]}{n} \geq \frac{h(U^n)}{n} - \Phi(D) - O\left(\frac{\log n}{n}\right).$$

Insight: The broader class of one-to-one codes allows for a potential rate reduction of $O(\log n/n)$ for finite n. This aligns with known results in lossless coding and highlights the precision of the framework.

For D-semifaithful codes, we redefine

$$Z_n(\alpha,\beta) = \int_{\mathcal{U}^n} \exp_2 \left\{-\alpha L[\phi_n(u^n)] - \beta W[nD - \rho(u^n - \psi_n(\phi_n(u^n)))]\right\} \mathrm{d}u^n$$

where $W[x] = 0$ for $x \geq 0$ and $W[x] = \infty$ for $x < 0$. Here, $Z_n(\alpha,\beta) \leq \mathrm{Vol}\{\boldsymbol{z}:\ \rho(\boldsymbol{z}) \leq nD\} = \mathrm{Vol}\{\mathcal{S}_n(D)\} \leq 2^{n\Phi(D)}$ by ordinary Chernoff bounding. But using saddle-point integration:

$$\begin{aligned}
\mathrm{Vol}\{\mathcal{S}_n(D)\} &= \int_{\mathbb{R}^n} u(nD - \rho(\boldsymbol{z}))\mathrm{d}\boldsymbol{z} \\
&= \int_{\mathbb{R}^n} \frac{\mathrm{d}\boldsymbol{z}}{2\pi i} \int_{\mathrm{Re}\{s\}=c} \frac{e^{s(nD-\rho(\boldsymbol{z}))}\mathrm{d}s}{s} \\
&= \frac{1}{2\pi i} \int_{\mathrm{Re}\{s\}=c} \frac{\mathrm{d}s}{s} \exp\left\{n\left[sD + \ln\left(\int_{\mathbb{R}} e^{-s\rho(z)}\mathrm{d}z\right)\right]\right\} \\
&= \frac{2^{n\Phi(D)}}{s_\star\sqrt{2\pi n|f''(s_\star)|}} \cdot \left[1 + O\left(\frac{1}{n}\right)\right].
\end{aligned}$$

$$\frac{\mathbf{E}\{L[\phi_n(U^n)]\}}{n} \geq \frac{h(U^n) - \log \mathrm{Vol}\{\mathcal{S}_n(D)\}}{n} = \frac{h(U^n)}{n} - \Phi(D) + \frac{\log n}{2n} + O\left(\frac{1}{n}\right).$$

Application 2: A Tighter Bound for D-semifaithful Codes

Scenario: We impose a stricter pointwise distortion constraint, $\rho(\mathbf{u}^n - v^n) \leq nD$ for all $\mathbf{u}^n$, not just on average. This defines a D-semifaithful code.

Method: The analysis relies on a different extended Kraft inequality (Lemma 3) and a more precise estimation of the resulting volume term using the saddle-point method.

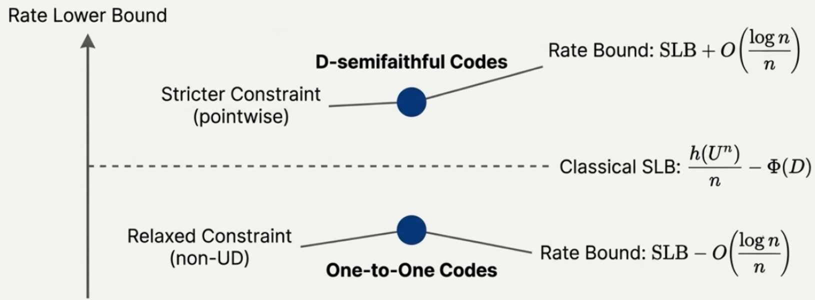
Result: The stricter constraint results in a higher rate lower bound, adding a positive redundancy term to the classical SLB.

$$\frac{\mathbb{E}[L]}{n} \geq \frac{h(U^n)}{n} - \Phi(D) + \frac{k' \log n}{2n} + O\left(\frac{1}{n}\right).$$

Where k' is the "effective number" of active distortion constraints.

Insight: The pointwise constraint is more demanding, and the framework precisely quantifies this cost as an additional $O(\log n/n)$ rate term.

A Unified View of Finite-Blocklength Refinements



The extended Kraft inequality framework provides a unified lens to analyze how different coding constraints modify the rate-distortion tradeoff for finite blocklengths, yielding both tighter and looser bounds relative to the classical SLB.

Application 3: Generalizing to Sliding-Window Distortion

Scenario: Consider distortion measures that depend on m consecutive symbols of the error signal, z^n. This allows for shaping the error's spectral properties or memory.

Example: Constraining both mean-squared error ($\sum z_t^2$) and first-lag autocorrelation ($\sum z_t z_{t-1}$).

Distortion Form:

$$\sum_{t=m \text{ to } n} \rho_j(z_t, \ldots, z_{t-m+1}) \leq nD_j$$

Method: The framework extends naturally. The integral bound now involves the dominant eigenvalue (spectral radius), $\lambda(\beta)$, of the sliding-window integral operator.

Asymptotic Result:

$$\liminf \frac{\mathrm{E}[L]}{n} \geq \bar{h}(U_\infty) - \inf_{\beta \in [0,\infty)^k} [\log \lambda(\beta) + \beta \cdot D]$$

Insight: The framework's core structure remains, adapting to incorporate memory by replacing a simple integral with a spectral radius calculation.

The Scenario

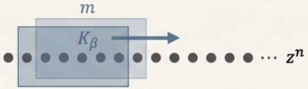
In many applications (e.g., image and audio), the perceived distortion depends not just on per-sample error, but on the structure and memory of the error signal. This can be captured by sliding-window distortion functions.

$$\rho(z^n) = \sum_{i=m}^{n} \rho(z_{i-m+1}^{i})$$

Example: Constraining both mean-square error ($\rho(z_i) = z_i^2$) and first-lag autocorrelation ($\rho(z_i, z_{i-1}) = -z_i * z_{i-1}$).

The Twist

The multi-dimensional integral in the Kraft inequality bound $\int \exp_2\{-\beta \cdot \sum \rho(\ldots)\}\, dz^n$ can be analyzed as an iterated application of an operator whose kernel is $K_\beta(\ldots) = \exp_2\{-\beta \cdot \rho(\ldots)\}$.



Example: Quantifying the Cost of an Autocorrelation Constraint

Problem Setup: Consider two constraints:

1. Quadratic Distortion: $\frac{1}{n}\sum z_t^2 \leq D$
2. First-Lag Autocorrelation: $\frac{1}{n}\sum z_t z_{t-1} \geq \theta$ (rewritten as $\frac{1}{n}\sum(-z_t z_{t-1}) \leq -\theta$)

Classical SLB (Constraint 1 only)

$$R(D) \geq \bar{h}(U_\infty) - \frac{1}{2}\log(2\pi e D)$$

Generalized SLB (Both Constraints)
Applying the spectral radius formulation yields a new, higher bound:

$$R(D, \theta) \geq \bar{h}(U_\infty) - \frac{1}{2}\log[2\pi e D(1-\theta^2)]$$

Resulting Rate Penalty:
The additional rate required to enforce the autocorrelation constraint is:

$$\Delta R = \frac{1}{2}\log\left(\frac{1}{1-\theta^2}\right)$$

Application 4: An Individual-Sequence Counterpart of the SLB

Scenario: Can we establish a rate-distortion bound for a single, deterministic source sequence u^n, without relying on an underlying probability distribution?

Method: 1. The model assumes the source is quantized, and the resulting reproduction sequence is compressed by a finite-state (FS) encoder.

2. A generalized Kraft inequality for information-lossless FS encoders is developed (Lemma 5).

The Key Substitution

Source Entropy

$$\frac{h(U^n)}{n}$$

is replaced by...

Normalized Lempel-Ziv (LZ) Complexity

$$\frac{c(u^n)\log c(u^n)}{n}$$

$c(u^n)$ is the number of phrases in the LZ parsing of the sequence.

Result: The derivation follows a similar path, but the probabilistic source entropy $h(U^n)/n$ is naturally replaced by a term from universal compression theory.

Insight: This result forges a fundamental link between rate-distortion theory and the algorithmic complexity of individual sequences.

The SLB for Individual Sequences

Main Result: For any information-lossless finite-state encoder, the compression rate is lower-bounded by:

$$\frac{1}{n}\sum \mathcal{L}[f(\sigma_t, v_t)] \geq \frac{c(u^n)\log c(u^n)}{n} - \Phi\left(\frac{1}{n}\sum \rho(u_t - v_t)\right)$$

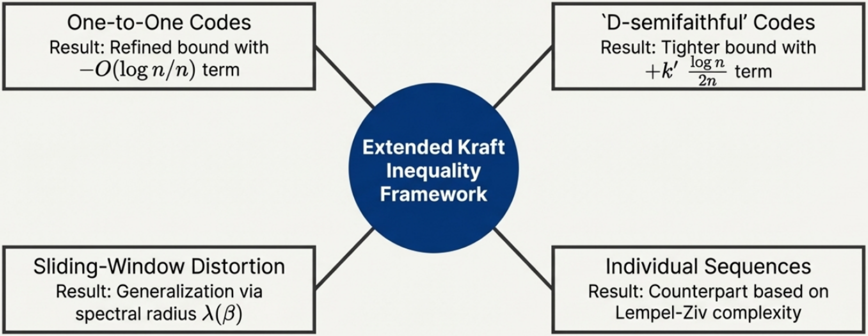
(Ignoring vanishing terms for large n)

| Probabilistic Setting |	| Individual Sequence Setting |
$R(D)$ (Expected Rate)	$\left(\frac{1}{n}\right)\sum \mathcal{L}$ (Actual Rate)
$h(U^n)/n$ (Entropy Rate)	$c(u^n)\log c(u^n)/n$ (LZ Complexity)
D (Expected Distortion)	$\left(\frac{1}{n}\right)\sum \rho$ (Actual Distortion)

Conclusion:

The SLB's fundamental structure holds even for individual sequences, with source entropy replaced by LZ complexity.

A Unifying Framework for Rate-Distortion Bounds



A single, powerful proof technique based on an extended Kraft inequality serves as a versatile tool. It not only recovers the classical SLB but also generates a family of refined and generalized bounds applicable to a wider range of coding constraints and source models.

Contribution and Significance

- This work demonstrates the power of a methodological framework rooted in the Kraft inequality for advancing rate-distortion theory.
- It provides sharper, non-asymptotic bounds for codes with specific structural properties (one-to-one, pointwise constraints).
- It extends the theory to handle more complex, structured distortion measures relevant to modern applications.
- Finally, it bridges the gap between the probabilistic and deterministic viewpoints by establishing an individual-sequence counterpart to the Shannon Lower Bound based on algorithmic complexity.

Collectively, these results deepen the understanding of the fundamental limits of lossy data compression.