Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Entropy Change in Sampling Predicts Downstream Performance in Neural Networks

**Jun-Jie Zhang**

zjacob@mail.ustc.edu.cn

Northwest Institute of Nuclear Technology

**Xiu-Cheng Wang**

Xidian University

**Nan Cheng**

Xidian University

**Long-Gang Pang**

Central China Normal University

**Tai-Jiao Du**

Northwest Institute of Nuclear Technology

**Deyu Meng**

Xi'an Jiaotong University

**Article**

**Keywords:** entropy auditing, Husimi spectrogram, subsampling, acquisition design, vision, MIMO, MRI

**Posted Date:** November 18th, 2025

**DOI:** https://doi.org/10.21203/rs.3.rs-8034605/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** There is **NO** Competing Interest.

# Entropy Change in Sampling Predicts Downstream Performance in Neural Networks

Xiu-Cheng Wang[2,3†], Jun-Jie Zhang[1*†], Nan Cheng[2,3], Long-Gang Pang[4], Taijiao Du[1], Deyu Meng[5,6*]

[1]Northwest Institute of Nuclear Technology, No. 28 Pingyu Road, Xi'an, 710024, Shaanxi, China.
[2]School of Telecommunications Engineering, Xidian University, No. 2 South Taibai Road, Xi'an, 710071, Shaanxi, China.
[3]State Key Laboratory of ISN, No. 2 South Taibai Road, Xi'an, 710071, Shaanxi, China.
[4]Key Laboratory of Quark and Lepton Physics (MOE) & Institute of Particle Physics, Central China Normal University, No. 152 Luoyu Road, Wuhan, 430079, Hubei, China.
[5]Ministry of Education Key Lab of Intelligent Networks and Network Security, Xi'an Jiaotong University, No. 28 Xianning West Road, Xi'an, 710049, Shaanxi, China.
[6]School of Mathematics and Statistics, Xi'an Jiaotong University, No. 28 Xianning West Road, Xi'an, 710049, Shaanxi, China.

*Corresponding author(s). E-mail(s): zjacob@mail.ustc.edu.cn; dymeng@mail.xjtu.edu.cn;
Contributing authors: xcwang_1@stu.xidian.edu.cn; dr.nan.cheng@ieee.org; lgpang@ccnu.edu.cn; dutaijiao@nint.ac.cn;
†Both authors contributed equally to this work.

**Abstract**

Data acquisition, the physical interface between the world and learning systems, fundamentally determines how much information is available before any model is trained. Existing evaluation and optimization methods typically rely on learned losses or domain-specific heuristics. Here, we introduce a training-free scalar—band-entropy change—which quantifies how an acquisition process disturbs the spectral structure of the signal. Across three domains—vision (patch masking), wireless multiple-input multiple-output (MIMO) systems (pilot/antenna subsampling), and magnetic resonance imaging (MRI) (***k***-space undersampling)—we show experimentally that the magnitude of band-entropy change, computed from raw measurements, provides a useful, training-free indicator of downstream performance: smaller values are consistently associated with higher classification accuracy or reconstruction quality. Our results motivate a task-agnostic, instrument-aware diagnostic—entropy auditing—that can evaluate or guide acquisition choices prior to training. This framework bridges the gap between physical measurement principles and machine learning, offering a physics-informed approach to optimizing sampling strategies and predicting downstream performance across various domains.

**Keywords:** entropy auditing, Husimi spectrogram, subsampling, acquisition design, vision, MIMO, MRI

## Introduction

Data acquisition—the physical interface between the world and a learning system—fundamentally determines what can be learned [1–4]. In various scientific and engineering fields, measurements are inherently constrained by factors such as time, energy, bandwidth, and hardware limitations.

This makes subsampling—or reducing the amount of data collected—necessary in many applications, including accelerated magnetic resonance imaging (MRI) [5, 6], computed tomography (CT), cryo-electron microscopy [7], massive multiple-input multiple-output (MIMO) systems [8–12], and modern vision models [13–17]. A common empirical observation across these domains is that, under the same sampling budget, randomized or variable-density acquisition strategies often outperform strictly periodic or regular sampling patterns. While this observation has been widely documented, the physical reasons behind it have remained unclear.

Existing theoretical frameworks each explain part of the puzzle. Compressed sensing and incoherence theory guarantee signal recovery under sparsity constraints [18–23]. Information-theoretic and optimal-design approaches quantify expected utility through mutual or Fisher information [24–28]. More recently, data-centric paradigms estimate sample importance or transferability after training [29–35]. While valuable, these methods rely on task-specific priors, learned losses, or model dynamics, and thus provide limited insight into the information structure of the measurements themselves. A training-free, instrument-aware criterion that directly links acquisition geometry to information preservation would therefore be broadly useful.

Here, we provide a complementary aspect: instead of predicting performance through a model's behavior, we can analyze the data's information content directly, before any learning takes place. We view acquisition as a perturbation of a signal's local spectrum on the Nyquist torus [36–40]. We quantify this perturbation using the Shannon entropy of the Husimi distribution [41–43], defining the *band-entropy change* $\Delta S_{\mathcal{B}}$. Beyond serving as an estimator, the Husimi transform also acts as a *bridge between measurement and representation*: its spatial–frequency parameters $(\texttt{win}, \sigma, \texttt{hop})$ should be chosen to match both the instrument's effective resolution and the receptive-field scale of the downstream network, so that entropy auditing evaluates not only what is measured but also what the model can represent. This measure captures how the acquisition process redistributes spectral energy within the instrument's resolution. Our central proposition—the *Entropy-Preserving Acquisition (EPA)* principle—is that, sampling strategies that minimize the magnitude of spectral-entropy perturbation $|\Delta S_{\mathcal{B}}|$ tend to preserve the instrument-resolved information most effectively.

This principle explains two opposite but unified regimes. In spatial-domain masking (as in vision or wireless antenna selection), regular periodic patterns coherently fold frequencies, increasing spectral entropy; randomized masking breaks lattice coherence and preserves entropy on average. In frequency-domain undersampling (as in MRI), missing coefficients remove spectral mass and decrease entropy, but again, the acquisition that minimizes $|\Delta S_{\mathcal{B}}|$ best preserves recoverable structure. We provide analytical derivations (see Supplementary Information) showing that periodic subsampling increases $S_{\mathcal{B}}$ through coherent aliasing, while random patterns preserve it statistically. Empirically, $\Delta S_{\mathcal{B}}$ computed directly from raw data—without any learning—strongly correlates with downstream performance (Pearson $|r| \approx 0.9$ across domains). This behavior aligns with long-standing empirical results on random and blue-noise sampling [44–46], providing a unified, information-theoretic rationale.

By grounding data-centric learning in the physics of information preservation, the EPA framework elevates acquisition from empirical tuning to a measurable, theoretically interpretable process. It offers a domain-agnostic, training-free diagnostic—*entropy auditing*—for evaluating and guiding sampling, pilot, and sensor design prior to training, bridging the geometry of measurement with the learnability of representations.

# 1 Results

We first present empirical results showing that a simple, training-free metric—the band-entropy change—predicts downstream performance across vision, wireless, and MRI. Subsection 1.4 then develops the minimal phase-space framework that explains these observations.

## 1.1 Visual recognition: entropy auditing predicts task difficulty

To test the predictive power of the EPA principle, we evaluated a visual recognition task using a Vision Transformer (ViT) [16] trained on mini-ImageNet [47]. We compared two patch-subsampling strategies—**Uniform** (periodic lattice) and **Random** (uniform random selection without replacement)—across matched observation budgets corresponding to sampling intervals $k = 2, 4, 8$.

Random masking yielded higher test accuracy than Uniform masking at all sampling intervals (Fig. 1a). Accuracy shows a strong monotonic association with the mean $\Delta S_{\mathcal{B}}$ (smaller $\Delta S_{\mathcal{B}}$, higher

accuracy). Pearson coefficients estimated from four budget points per method indicate a large-magnitude correlation but wide confidence intervals; we therefore treat this as a qualitative trend rather than a precise effect size.

The distribution of $\Delta S_{\mathcal{B}}$ (Fig. 1b) provides further mechanistic insight. The violin plots reveal unimodal distributions for all conditions. Two key trends are evident: first, for any given $k$, the distribution for Uniform masking is systematically shifted towards higher $\Delta S_{\mathcal{B}}$ values compared to Random masking. Second, as the sampling interval $k$ increases (sparsity increases), the entire distribution of $\Delta S_{\mathcal{B}}$ shifts to the right for both methods, indicating a greater perturbation of the signal's spectral structure. This systematic increase in spectral disorder underlies the corresponding drop in classification accuracy.

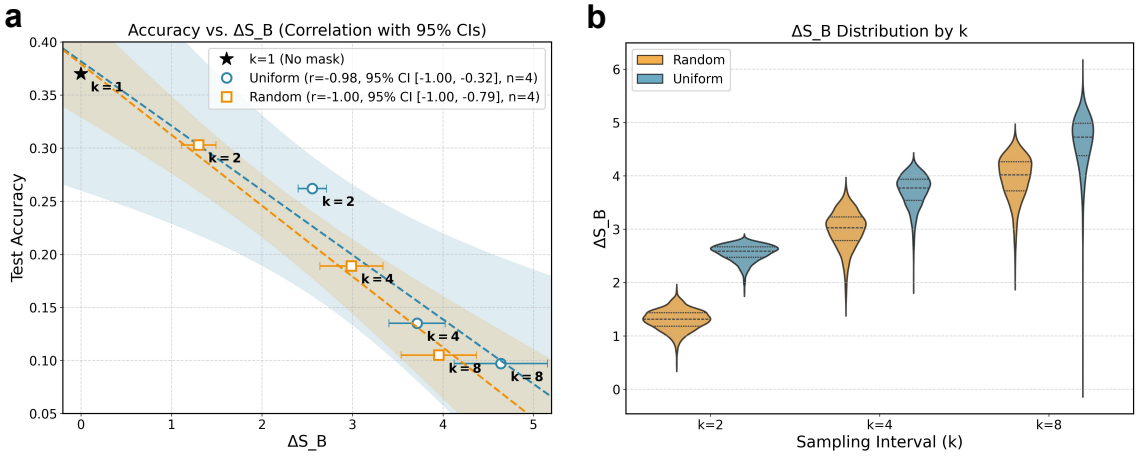



**Fig. 1 Entropy auditing predicts visual recognition performance. (a)** Test accuracy versus the mean band-entropy change ($\Delta S_{\mathcal{B}}$) for different sampling strategies and intervals. Each point represents the mean $\Delta S_{\mathcal{B}}$ (x-axis) and the corresponding test accuracy (y-axis) for one method and $k$ value; the star denotes the baseline ($k$=1, no masking). Dashed lines show ordinary-least-squares (OLS) fits with shaded 95% confidence intervals for the mean regression line, revealing strong negative correlations (empirical slopes are large in the sampled budgets). Pearson $r$ values are reported with 95% confidence intervals estimated by Fisher's z-transform; because the number of budget points per method is small ($n$=4), these correlations should be interpreted as qualitative trend evidence rather than precise effect-size estimates. **(b)** Distribution of $\Delta S_{\mathcal{B}}$ values for different methods and sampling intervals $k$. The violin plots show the kernel density estimate, revealing unimodal distributions that are typically broader in the body and taper at the ends. The internal dashed lines mark the quartiles (25th, 50th, 75th percentiles). The distribution for Uniform masking is consistently right-shifted compared to Random masking at every $k$, and the overall distribution of $\Delta S_{\mathcal{B}}$ shifts to higher values as $k$ increases.

## 1.2 MIMO: Entropy auditing explains why randomized pilot and antenna subsampling performs better

Accurate channel state information (CSI) is crucial for the performance of multi-antenna communication systems, yet the pilot overhead required for its estimation grows rapidly with array size. When the number of antennas becomes large, the system must subsample pilots or deactivate antennas, inevitably losing part of the spatial information. We examine how different subsampling strategies influence the recoverability of CSI and how entropy auditing can predict this effect before any reconstruction or learning.

We consider a narrowband MIMO system, where the channel matrix $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ describes the complex gain between $N_t$ transmit and $N_r$ receive antennas. Under a fixed pilot budget, only the entries indexed by $\Omega \subseteq \{1, \ldots, N_r\} \times \{1, \ldots, N_t\}$ are observed, and the reconstruction algorithm estimates the full $\widehat{\mathbf{H}}$ from $\mathbf{H}_\Omega$. Two practically feasible sampling patterns are compared under exactly matched observation budgets:

- **Uniform (periodic lattice):** antennas are deactivated periodically every $d$ positions along both transmit and receive axes, producing a regular grid of observations.
- **Random (uniform without replacement):** the same number of antenna indices are selected randomly along each axis, removing long-range periodicity while maintaining the same pilot cost.

This setting mirrors the spatial masking problem in vision tasks: turning off antennas corresponds to masking rows and columns of $\mathbf{H}$. From the entropy-auditing perspective, periodic deactivation

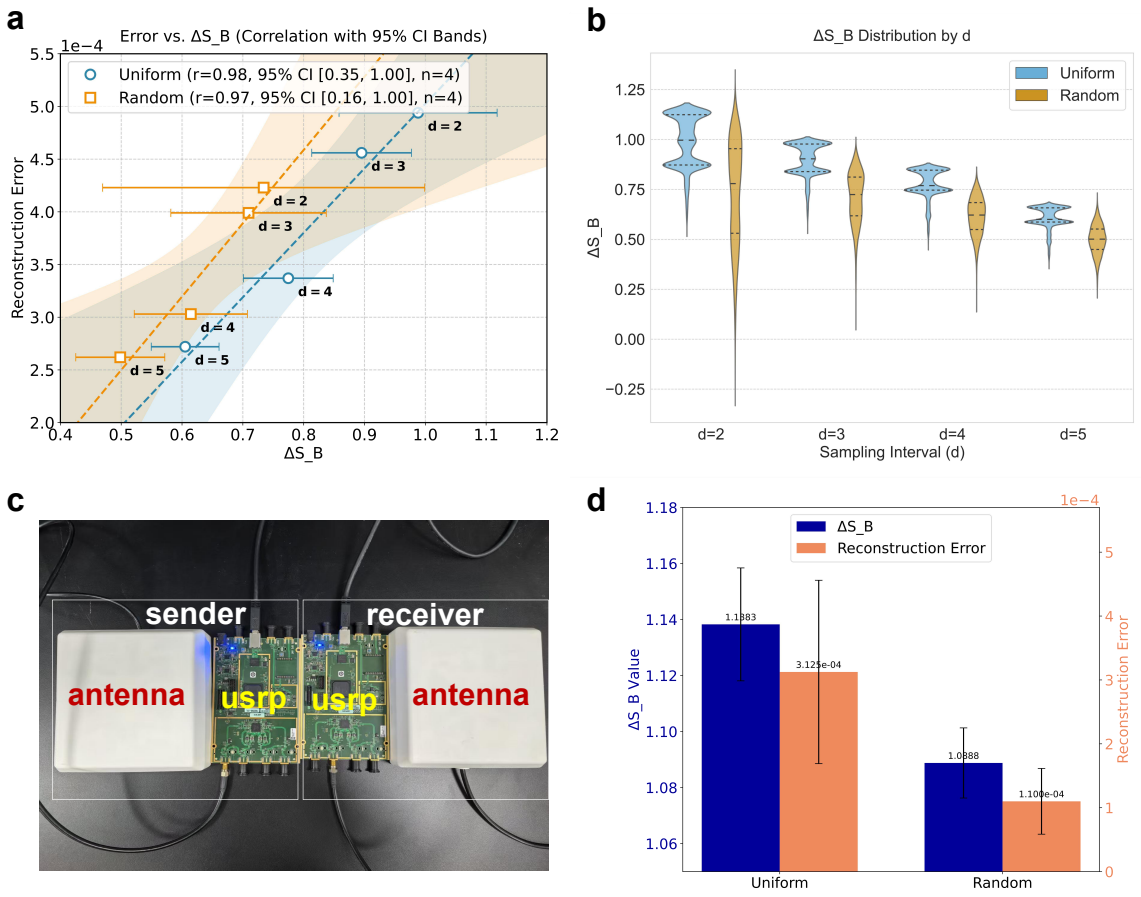



**Fig. 2 Randomized subsampling preserves spectral entropy and improves CSI recovery. (a)** Relationship between normalized mean squared error (NMSE, lower is better) and mean band-entropy change $\Delta S_{\mathcal{B}}$ for Uniform and Random antenna deactivation under matched pilot budgets determined by interval $d$. Dashed lines indicate OLS fits; shaded regions show 95% confidence intervals for the mean regression line (as in Fig. 1). A strong positive correlation is observed for both strategies (Pearson $r \approx 0.97$): larger entropy perturbation predicts higher NMSE. **(b)** Distributions of $\Delta S_{\mathcal{B}}$ for varying $d$. For all sampling intervals, the Uniform scheme shows higher $\Delta S_{\mathcal{B}}$ values than Random, and both distributions shift toward smaller $\Delta S_{\mathcal{B}}$ as $d$ increases, indicating less spectral perturbation at denser sampling. **(c)** Over-the-air (OTA) validation using an 8×8 uniform planar array with software-defined radios (SDRs) and patch antennas. **(d)** Experimental results under identical pilot budgets. The Random scheme achieves both lower NMSE and smaller $|\Delta S_{\mathcal{B}}|$ than Uniform. Error bars denote across-trial standard deviation.

leads to lattice-coherent folding of spectral energy and increases the disorder of the Husimi density, whereas random deactivation breaks the lattice coherence and tends to preserve entropy in expectation. Consequently, the band-entropy change $\Delta S_{\mathcal{B}}$ should reflect the intrinsic difficulty of CSI recovery.

Figure 2 summarizes the simulation and OTA experiments under matched pilot budgets. In all tested configurations, randomized subsampling consistently outperforms the periodic lattice pattern, and the performance gap widens as the channel is more sparsely sampled. Across budgets, NMSE increases as the mean $\Delta S_{\mathcal{B}}$ increases (Fig. 2a), with large-magnitude Pearson correlations estimated from a small number of budget points; we report them as trend indicators. The violin plots in Fig. 2b show unimodal distributions of $\Delta S_{\mathcal{B}}$: the periodic scheme always exhibits higher entropy change, while increasing sampling density (larger $d$) reduces overall entropy perturbation. These trends remain consistent in the OTA validation (Fig. 2c–d). Using the same pilot budget, the random subsampling achieves about one-third lower NMSE and a smaller $|\Delta S_{\mathcal{B}}|$ compared with the periodic counterpart. The ranking remains unchanged when the random mask is either fixed for all coherence intervals or refreshed for each one, indicating robustness to implementation details.

### 1.3 MRI: entropy auditing predicts reconstruction quality under $k$-space undersampling

We next evaluate EPA in a complementary frequency-domain setting, where entropy decreases rather than increases. We apply it to accelerated multi-coil magnetic resonance imaging (MRI), a domain where the underlying physics of acquisition differs from the previous examples. In MRI, subsampling occurs in $k$-space (the frequency domain), not the spatial domain [5]. This distinction provides a complementary test for our core hypothesis: if the band-entropy change $\Delta S_{\mathcal{B}}$ is truly a measure of structural information change, it should predict reconstruction difficulty regardless of whether the entropy increases or decreases.

We use different undersampling masks to fully-sampled $k$-space data to emulate accelerated acquisition. Missing coefficients are zero-filled, producing images via inverse fast Fourier transform (FFT). These inputs are reconstructed by a U-Net [48], while $\Delta S_{\mathcal{B}}$ is computed directly on the inputs relative to the ground truth, without any training. We compare three canonical mask types—Uniform, Random, and Poisson—under multiple acceleration factors.

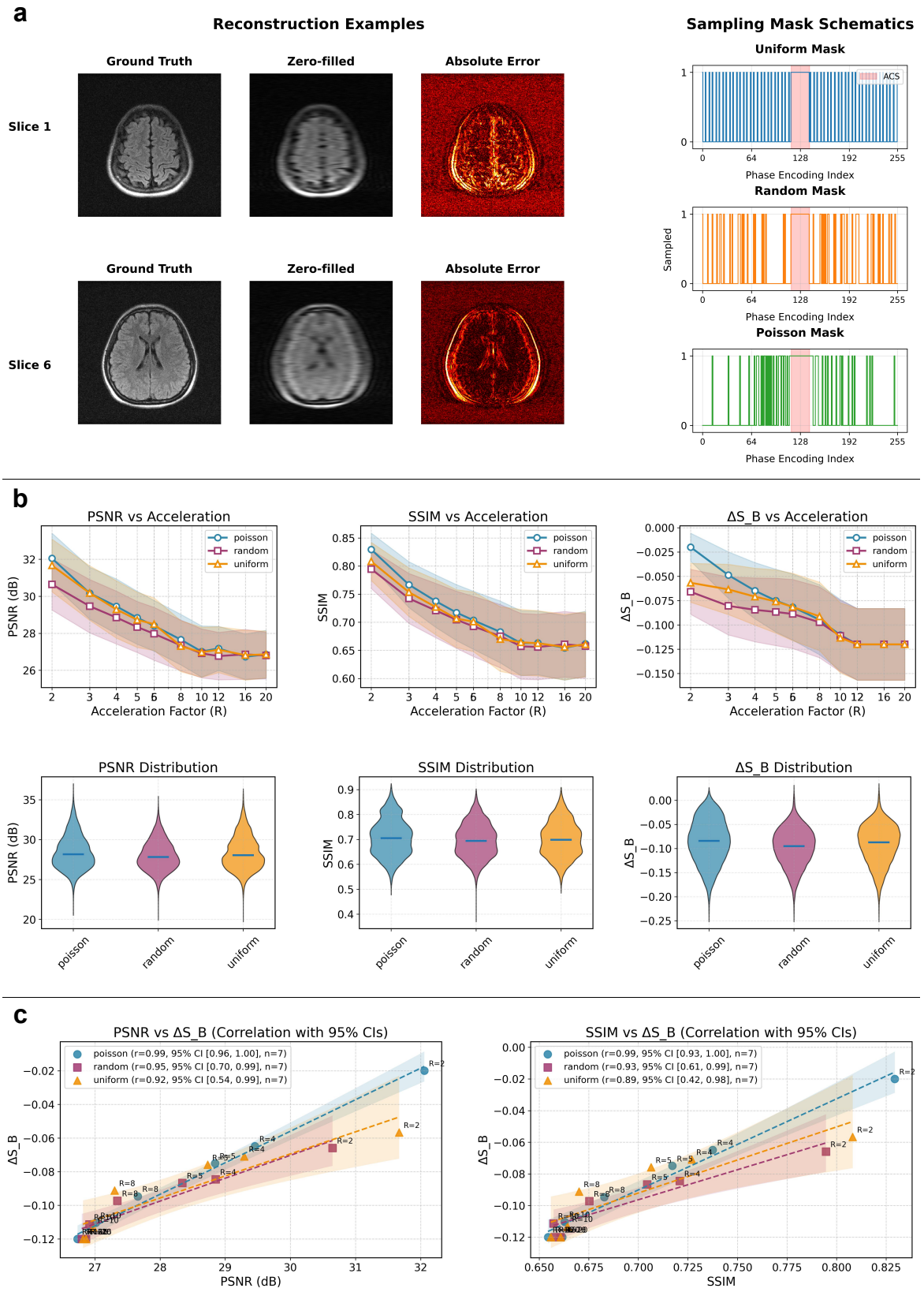



**Fig. 3 Band-entropy change $\Delta S_{\mathcal{B}}$ predicts reconstruction quality in accelerated MRI. (a)** Reconstruction examples and mask schematics. The right panel illustrates 1D $k$-space masks. The acquisition process involves multiplying the full $k$-space by a binary mask, which directly removes frequency data. **(b)** Performance trends versus acceleration factor ($R$). Top: Reconstruction quality (peak signal-to-noise ratio, PSNR; structural similarity index, SSIM) and band-entropy change for different mask methods. Bottom: The corresponding distribution of the quality and $\Delta S_{\mathcal{B}}$ values. The empirical performance hierarchy (Poisson > Uniform > Random) is consistently mirrored by the entropy change. Strategies that cause the smallest perturbation to the original entropy (i.e., $\Delta S_{\mathcal{B}}$ is closest to zero) consistently perform best. **(c)** Quality–entropy association. Scatter plots show the correlation between mean $\Delta S_{\mathcal{B}}$ and mean PSNR/SSIM. Dashed lines indicate OLS fits; shaded regions show 95% confidence intervals for the mean regression line (as in Fig. 1).

The results are summarized in Fig. 3. Unlike spatial masking, which causes frequency folding and leads to a more uniform, higher-entropy mixture of spectral components ($\Delta S_{\mathcal{B}} > 0$), $k$-space undersampling directly removes frequency components. The act of setting parts of the $k$-space to zero constitutes a direct loss of spectral information. A sparser, less complex spectrum has a lower Shannon entropy. Consequently, for all MRI sampling strategies, the band-entropy of the aliased input is lower than the ground truth, resulting in a consistently negative band-entropy change ($\Delta S_{\mathcal{B}} < 0$).

Within this physical regime, the proposed framework holds robustly: the magnitude of the entropy change, $|\Delta S_{\mathcal{B}}|$, serves as a proxy for the difficulty of the downstream task. An acquisition strategy that minimally perturbs the original signal's entropy (i.e., results in a $\Delta S_{\mathcal{B}}$ closest to zero) best preserves the recoverable structure.

Fig. 3b quantifies this relationship. The empirical performance ranking, measured by Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), is consistently Poisson > Uniform > Random. This ordering was reflected by the computed $\Delta S_{\mathcal{B}}$ values under the chosen Husimi parameters and reconstruction network. The purely Random mask, by indiscriminately removing frequency coefficients across the spectrum, causes the most severe and unstructured information loss. The Uniform mask, while inducing coherent aliasing artifacts (Fig. 3a), preserves a structured subset of frequencies, which proves more beneficial for the convolutional nature of the U-Net than the chaotic removal by the random mask. The Poisson mask performs best because its variable-density nature is physically informed: it preserves the crucial, high-energy low-frequency components (essential for image contrast and structure) while randomly sampling the higher frequencies (related to finer details), striking an optimal balance between acceleration and information preservation.

Meanwhile, Fig. 3c confirms the positive correlation between reconstruction quality and $\Delta S_{\mathcal{B}}$. A $\Delta S_{\mathcal{B}}$ value closer to zero is robustly associated with higher PSNR and SSIM, demonstrating that the change in band-entropy, irrespective of its sign, serves as a reliable, task-agnostic proxy for task difficulty.

## 1.4 Theoretical Framework: A Phase-Space Foundation for Physics-Informed Acquisition

The consistent empirical findings across vision, wireless, and MRI domains reveal a fundamental regularity: acquisition strategies that minimize spectral entropy perturbation consistently yield optimal downstream performance. This emergent principle motivates a unified theoretical framework grounded in phase-space physics, connecting classical signal processing with modern machine learning through information-theoretic foundations. Here, we outline the framework and state two general guarantees; full derivations, constants, and boundary terms are in the *Supplementary Information*.

### *Bridging Physical Measurement and Learned Representation with Phase-Space Representation*

Inspired by methods from quantum mechanics and time-frequency analysis, we model signals through their phase-space representation—a joint space-frequency description that naturally captures the information content accessible through physical measurement. For a signal $I(\mathbf{x})$, we employ the Husimi distribution [37], a smoothed Wigner transform that provides a nonnegative, instrument-aligned spectrogram:

$$\rho_I(\mathbf{x};\mathbf{k}) \;=\; (\Phi * W_I)(\mathbf{x};\mathbf{k}). \tag{1}$$

where $W_I$ is the Wigner distribution, $\Phi$ represents the finite resolution kernel of the measurement instrument. Here "$*$" denotes convolution in full phase space $(\mathbf{x},\mathbf{k})$. This formulation answers a physically fundamental question: "At spatial location $\mathbf{x}$, how is the signal's energy distributed over frequencies $\mathbf{k}$ given the instrument's resolution limits?"

Practically, this implies a dual-alignment rule: the kernel $\Phi$ (realized by $(\texttt{win},\sigma,\texttt{hop})$ in our Husimi implementation) should align with both the sensor's space–frequency resolution and the network's receptive-field hierarchy; otherwise $\Delta S_{\mathcal{B}}$ may reflect a scale mismatch rather than intrinsic information change.

### *Quantifying Information Preservation using Spectral Entropy*

To quantify how acquisition perturbs a signal's fundamental information structure, we define the normalized local spectrum:

$$\mathcal{N}_{\mathcal{B}}[\rho_I]\,(\mathbf{x};\mathbf{k}) \;=\; \frac{\rho_I(\mathbf{x};\mathbf{k})}{\int_{\mathcal{B}} \rho_I(\mathbf{x};\tilde{\mathbf{k}})\,\tilde{\mathbf{k}}}. \tag{2}$$

which forms a probability distribution over frequencies at each spatial location. The local spectral entropy:

$$s_{\mathcal{B}}(\mathbf{x}) \;=\; -\int_{\mathcal{B}} \mathcal{N}_{\mathcal{B}}[\rho_I]\,(\mathbf{x};\mathbf{k})\,\log \mathcal{N}_{\mathcal{B}}[\rho_I]\,(\mathbf{x};\mathbf{k})\,\mathbf{k}. \tag{3}$$

measures the "disorder" or "spread" of frequency content locally, while the global band-entropy $S_{\mathcal{B}}$ averages this measure across the entire field.

This entropy formulation captures a deep physical intuition: signals with concentrated spectral energy (e.g., smooth images with low complexity) exhibit low entropy, while those with dispersed spectra (e.g., textured or noisy signals) show high entropy. Crucially, this measure is computable directly from physical measurements without requiring task-specific training.

***Entropy Preservation as a Physical Principle for Task Difficulty.***

From a physical standpoint, the entropy of a signal's local spectrum quantifies how evenly its energy is distributed across accessible frequency modes. Acquisition, regardless of its implementation, acts as a *perturbation* to this spectral energy landscape. If the acquisition process redistributes energy coherently (as in periodic sampling) or removes it selectively (as in undersampling), the phase-space disorder—and hence the spectral Shannon entropy—is altered. Because the entropy of the Husimi distribution reflects the number of effectively accessible degrees of freedom of the measured field, a perturbation in entropy directly corresponds to a change in the amount of physically recoverable information. Preserving spectral entropy therefore means preserving the fundamental information capacity of the signal as seen by any downstream model.

Formally, our analysis yields a physics-grounded design principle for optimal acquisition: *the optimal sampling strategy minimizes the magnitude of the spectral-entropy perturbation*

$$\Delta S_{\mathcal{B}} \;=\; S_{\mathcal{B}}[\text{acquired}] \;-\; S_{\mathcal{B}}[\text{reference}]. \tag{4}$$

We refer to this procedure as *entropy auditing*. In other words, an acquisition that leaves the entropy of the local spectrum invariant also leaves the information structure of the field intact, resulting in easier learning or reconstruction tasks.

This principle naturally explains our empirical findings across domains and **can be formally justified within our phase-space framework** (see Supplementary Information), where explicit kernel formulations for Uniform and Random masks analytically demonstrate that periodic sampling increases $\Delta S_{\mathcal{B}}$ through coherent aliasing, whereas random sampling statistically preserves it in expectation.

By analyzing the entropy of the resulting mixed spectra, we show:

- In spatial masking (vision and wireless), periodic patterns introduce coherent aliasing that *increases* entropy ($\Delta S_{\mathcal{B}} > 0$);
- In frequency-domain sampling (MRI), missing coefficients remove spectral mass and *decrease* entropy ($\Delta S_{\mathcal{B}} < 0$);
- In both cases, the best performance is obtained when $|\Delta S_{\mathcal{B}}|$ is smallest, i.e., when the original spectral organization is least perturbed.

Thus, entropy preservation arises as a physics informed information-theoretic criterion, providing a direct, training-free measure of task difficulty and a principled bridge between measurement geometry and representational fidelity.

# 2 Methods

## 2.1 Visual Recognition on mini-ImageNet

***Dataset, Model architecture and training.***

We conducted experiments on the mini-ImageNet dataset, comprising 100 classes with 600 images per class (60,000 total images). All images were resized to $256 \times 256$ pixels and normalized using ImageNet statistics.

We employed a Vision Transformer (ViT-B/16) backbone, which processes $256 \times 256$ images as sequences of $16 \times 16$ non-overlapping patches (256 patches per image). The model was initialized with standard ViT weights and adapted for the 100-class classification task. For each experimental condition, separate models were trained from scratch using the AdamW optimizer (learning rate $10^{-3}$, weight decay 0.05) with a cosine learning-rate schedule. Training lasted 300 epochs with batch size 256.

***Masking strategies.***

We systematically compared patch-level subsampling strategies under matched observation budgets. For each sampling interval ($k = 2, 4, 8$), corresponding to keeping $1/k^2$ of the patches, we implemented:

- **Uniform:** Periodic sampling on a fixed $k \times k$ grid, retaining patches at regular intervals along both spatial dimensions.
- **Random:** Uniform random selection without replacement of the target number of patches.

To ensure fair comparison and eliminate variance from mask resampling, a single fixed random mask per budget was shared across all training and evaluation runs.

***Training protocol.***

All models were trained under identical conditions except for the masking strategy. The visible patches were processed by the standard ViT encoder, while masked patches were replaced with a learnable mask token. The classification head remained unchanged across conditions. We used cross-entropy loss and reported top-1 accuracy on the held-out test set. Results represent the mean and 95% confidence interval over three independent runs with different random seeds.

***Entropy auditing with ensemble averaging.***

The band-entropy change $\Delta S_{\mathcal{B}}$, a training-free diagnostic of information loss, was computed on the test images using a local Husimi-based procedure. For random masking, an ensemble averaging approach was employed: instead of averaging entropy values over multiple random mask instances, we first averaged the masked images across ten independent random realizations and then computed the entropy of this ensemble-averaged image. This procedure yields smoother and more stable entropy estimates by reducing high-frequency artifacts inherent to individual random mask instances.

The computation used a window size of 128 pixels with Gaussian smoothing parameter $\sigma = 64$ pixels and hop size of 32 pixels (50% overlap). These parameters were chosen to capture the dominant spatial–frequency structure of natural images at the patch level; in practice they represent heuristics that worked well for our data. A 128×128 window roughly covers an 8×8 patch region (for 16×16 patches on a 256×256 image), enabling the entropy estimate to reflect coarse structural organization rather than per-patch noise. The moderate Gaussian width ($\sigma = 64$) balances spatial and frequency resolution, while a 50% overlap provides stable averaging without excessive redundancy. As shown in Supplementary Information, the computed $\Delta S_{\mathcal{B}}$ depends on the Husimi parameters. For most scenarios, the results are robust across several parameters. However, for rare cases where the Husimi parameter is inconsistent with the feature size of the downstream network, we recommend reporting these parameters and testing robustness when applying the method to new sensors or datasets.

## 2.2 Wireless MIMO: simulation and over-the-air experiments

***Channel models and configurations.***

We consider narrowband MIMO channels represented by complex matrices $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$, where $N_t$ and $N_r$ denote the numbers of transmit and receive antennas. For simulation studies, $\mathbf{H}$ is generated using a clustered geometric channel model with Laplacian angular spreads. Each realization includes $L$ propagation paths characterized by complex gains, angles of arrival (AoA), and angles of departure (AoD), drawn from physically plausible distributions. Array configuration is kept fixed with 64 sender and 16 reciver with half-wavelength antenna spacing. Additive noise is introduced according to a predefined signal-to-noise ratio (SNR) setting of 15 dB.

***Pilot budgets and subsampling strategies.***

For each experiment, a fixed pilot budget defines the total number of observed entries in $\mathbf{H}$. Two acquisition schemes are evaluated under identical budgets:

- **Uniform:** antennas are deactivated periodically every $d$ positions along both transmit and receive axes, forming a deterministic grid pattern;
- **Random:** the same number of antenna indices are selected uniformly at random (without replacement) along each axis, removing long-range periodicity while maintaining the same sampling ratio.

The deactivation interval $d \in \{2, 3, 4, 5\}$ controls the subsampling rate, and all other conditions are held constant. Unless otherwise stated, the random mask is redrawn once per channel coherence interval; a fixed-mask variant reuses the same draw throughout the entire evaluation.

***Reconstruction and evaluation metrics.***

To isolate the effect of sampling geometry, all conditions employ the same reconstruction operator, we have the neural network directly output the reconstructed CSI matrix and use MSE to calculate the error statistics.

***Entropy auditing of channel structure.***

We compute Husimi-based local spectral entropy on the magnitude field $|\mathbf{H}|$. A two-dimensional Gaussian window of size $\texttt{win} = 4$ with standard deviation $\sigma = 1$ is applied, sliding with stride $\texttt{hop} = 1$ and zero padding at the borders to ensure full coverage. For each window, we take a 2D discrete Fourier transform with orthonormal scaling, compute the power spectrum on the discrete Fourier grid, and normalize it to a probability distribution by dividing by the total spectral energy. The Shannon entropy is then evaluated per window and averaged over all windows to obtain the global band-entropy $S_{\mathcal{B}}$. For each sample, the band-entropy change is defined as $\Delta S_{\mathcal{B}} = S_{\mathcal{B}}[\text{masked}] - S_{\mathcal{B}}[\text{original}]$, where the original is the fully sampled reference. This configuration resolves the sharp spatial transitions introduced by row/column deactivation while limiting spectral leakage via Gaussian tapering.

These parameters were selected empirically to match the spatial scale of antenna deactivation. A $4 \times 4$ window captures one to two neighboring rows or columns around a deactivated antenna, providing sufficient context to detect spectral spreading without averaging over the entire array. The Gaussian width $\sigma = 1$ (approximately one quarter of the window size) suppresses spectral leakage while preserving the sharp frequency transitions characteristic of lattice deactivation. The stride $\texttt{hop} = 1$ ensures that all boundary regions are analyzed without omission, which is essential for small matrices such as 8×8 or 16×64. Larger windows or strides were tested but yielded smoother entropy maps and weaker discrimination between periodic and random subsampling patterns.

***OTA validation.***

To corroborate the simulation results, we perform OTA measurements using an 8×8 uniform planar array equipped with patch antennas and two SDRs. The experiment is conducted in a quasi-static indoor environment with line-of-sight propagation. System calibration, including gain and phase alignment, is carried out before data acquisition. Pilot tones are sequentially transmitted according to the index set $\Omega$ and recorded to form $\mathbf{H}_\Omega$. Each configuration is repeated across random placements, and error bars in Fig. 2 denote the standard deviation across trials. The OTA results are consistent with the simulation trends we observed: in our testbed randomized subsampling produced smaller $|\Delta S_{\mathcal{B}}|$ and lower NMSE on average under matched pilot budgets. Because OTA experiments include environmental variability, these measurements are presented with trial-wise error bars to convey variability across placements and realizations.

***Implementation and reproducibility.***

All simulations were implemented in MATLAB and Python with NumPy and SciPy backends. OTA experiments used software-defined radio (SDR) model, e.g., Ettus USRP B210 synchronized via clock reference, e.g., GPSDO or 10 MHz reference. Detailed configuration files and scripts for generating channel models, applying masks, and computing Husimi entropy will be made available upon publication.

## 2.3 MRI Reconstruction and Entropy Auditing

***Data and preprocessing.***

Experiments were conducted on multi-coil MRI $k$-space data stored as complex-valued arrays. When necessary, dimensions were reordered so that the phase-encoding axis was the last dimension. Fully sampled $k$-space data were transformed into the image domain using a centered, orthonormal two-dimensional inverse Fourier transform. Multi-coil magnitude images were produced via root-sum-of-squares (RSS) combination. All images were normalized to $[0, 1]$ using the maximum magnitude of their fully sampled reference.

***$k$-space undersampling.***

To emulate accelerated acquisition, one-dimensional binary masks were generated along the phase-encoding (PE) axis and replicated across frequency lines. A fixed 24-line central auto-calibration region (ACS) was always retained. The remaining lines were selected according to one of three sampling distributions:

- **Uniform:** deterministic periodic spacing;
- **Random:** *uniform random selection without replacement* over non-ACS candidates to match the target sampling budget;
- **Poisson:** variable-density sampling with probability $p(d) \propto 1/(1+10d^2)$, where $d$ is the normalized distance from $k$-space center (ACS probability set to zero).

Acceleration ratios from $R$=1/20 to $R$=1/2 were evaluated. To ensure independence and reproducibility, each slice used a deterministically derived seed from a global seed when generating its mask.

***Reconstruction network and training.***

A conventional 2D U-Net with four encoding–decoding stages was trained independently for each mask–ratio configuration. The input was the zero-filled RSS image and the target was the fully sampled RSS image. Networks were optimized with Adam (learning rate $2 \times 10^{-4}$) using an L1 loss for 20 epochs with a batch size of 128. Reconstruction quality was assessed on the held-out test set using PSNR and SSIM (single-channel, data range = 1.0); per-slice metrics were averaged to report the final score. Detailed layer definitions and full training procedures are provided in the Code.

***Entropy auditing.***

For each test slice, we computed the Husimi-based band-entropy $S_{\mathcal{B}}$ of both the fully sampled and undersampled (zero-filled) magnitude images and defined

$$\Delta S_{\mathcal{B}} = S_{\mathcal{B}}[\mathrm{ZF}] - S_{\mathcal{B}}[\mathrm{Full}]. \tag{5}$$

Following a local band-normalization formulation, each image was divided into Gaussian-windowed patches of size $w \times w$ (window standard deviation $\sigma$); for each patch we computed the 2D FFT power spectrum, normalized it within the Nyquist band to a probability distribution, evaluated its Shannon entropy, and averaged the result across all patches.

***Choice of Husimi parameters.***

We set $(w, \sigma, \mathrm{hop}) = (24, 24, 24)$. For MRI magnitude images, spectral energy varies smoothly over tens of pixels and the effective spatial–frequency resolution is governed by the acquisition point-spread function. A 24-pixel window corresponds to roughly one quarter of the phase-encoding field of view (96–128 pixels in our data), which balances two competing effects: (i) overly small windows or $\sigma$ overweight high-frequency noise and inflate entropy estimates; (ii) overly large windows over-smooth spatial variation and reduce sensitivity to structured aliasing. Setting the hop equal to the window size avoids redundant overlaps and yields stable global averages.

# 3 Discussion

Our findings show that the band-entropy change, $\Delta S_{\mathcal{B}}$, is not merely a convenient indicator but a physically grounded descriptor of data structure in phase space. Importantly, the Husimi transform that underlies its computation acts as a bridge between measurement and representation: its spatial–frequency parameters must be consistent not only with the physical resolution of the acquisition process but also with the receptive-field scale of the downstream network. This dual alignment allows entropy auditing to evaluate how well the measured information matches the model's ability to represent it.

Unlike conventional analyses confined to a single domain, our approach evaluates entropy in the joint space of conjugate variables—spatial and spectral, or more generally, configuration and momentum. This phase-space perspective enables simultaneous consideration of a signal's organization and its spectral diversity, which mirrors how neural networks internally balance locality and frequency during representation learning. As shown in our prior work on the uncertainty principle of neural networks [49] and quantum-inspired vulnerability analysis [50], the interplay between conjugate representations governs both the efficiency and stability of learning. By auditing how acquisition perturbs this joint structure, $\Delta S_{\mathcal{B}}$ provides a scalar measure of information preservation that remains valid across domains and architectures.

***Complementarity with existing theory.***

Entropy auditing addresses a different but complementary question to classical mutual- or Fisher-information analyses. Whereas those approaches quantify the statistical value of measurements given

a probabilistic model or prior, EPA focuses on the *physical structure* of what is measured—the redistribution of spectral energy caused by the acquisition itself. It requires neither priors nor trained models and can therefore operate before any learning takes place. In this sense, $\Delta S_{\mathcal{B}}$ provides a fast, model-agnostic proxy for information preservation when task models are unavailable, and a geometric sanity check when they are.

***Relation to mutual and Fisher information.***

Mutual-information and optimal-design criteria formalize how much uncertainty is reduced once a prior over signals or parameters is specified. EPA instead measures how strongly the acquisition perturbs the entropy of the instrument-resolved field. When priors are approximately stationary and Gaussian, the two viewpoints converge: smaller $|\Delta S_{\mathcal{B}}|$ implies less information loss and thus higher expected MI or Fisher information. However, EPA remains meaningful even when such priors are unknown or unreliable, making it a practical complement rather than a replacement.

***Practical implications.***

Because $\Delta S_{\mathcal{B}}$ can be computed directly from raw measurements and proposed masks, it can enable rapid pre-training evaluation of candidate acquisition strategies. Practitioners can compare candidate sampling patterns by their $|\Delta S_{\mathcal{B}}|$ values on representative measurements—using identical Husimi parameters and preprocessing—and quickly eliminate designs that induce large entropy perturbations through coherent aliasing or excessive information removal. This simple diagnostic can reduce costly training cycles and provide physical intuition for why certain randomized or variable-density patterns generalize better across domains.

Moreover, when tuning the Husimi parameters, we should consider both the instrument's spatial–frequency resolution and the network's representational scale—the Husimi transform thereby serves as an explicit physical–algorithmic bridge linking data acquisition and model architecture.

***Limitations and outlook.***

Our approach is built on a few explicit assumptions (approximate local stationarity, sufficient spatial averaging by the smoothing kernel, and non-vanishing local band energy); these and the finite-$N$ concentration bounds (Assumptions A1–A6) are listed in the Supplementary Information. Moreover, in practice, the Husimi estimator — and therefore $\Delta S_{\mathcal{B}}$ — depends on choices such as window size, smoothing width and hop length, as well as the instrument's effective resolution and the measurement SNR; we therefore recommend routine robustness checks when applying the method to new sensors or datasets. Finally, because downstream performance depends on model class and training choices, $\Delta S_{\mathcal{B}}$ should be regarded as a practical, training-free diagnostic that complements — rather than replaces — prior- or model-based design criteria.

Looking forward, natural next steps are to quantify the estimator variance more precisely, extend the theory to nonstationary sensors and nonlinear measurement operators, and explore how entropy auditing might be incorporated as a regularizer or objective in learned acquisition schemes; more broadly, we view this work as one component of a physics-informed language for connecting measurement geometry and representation learning.

# 4 Conclusion

We introduced an instrument-aware perspective that connects acquisition geometry, spectral structure, and learnability through the band-entropy change $\Delta S_{\mathcal{B}}$. In our cross-domain experiments (vision, wireless, MRI) smaller $|\Delta S_{\mathcal{B}}|$ tended to align with better downstream performance, suggesting that $\Delta S_{\mathcal{B}}$ can serve as a compact, training-free diagnostic for comparing acquisition choices prior to learning. Entropy auditing is intended to complement classical information-theoretic and optimal-design criteria by providing a geometry- and measurement-focused viewpoint; further work is needed to integrate these perspectives under broader priors and operational constraints.

Beyond the specific domains studied here, entropy auditing establishes a physically interpretable foundation for data-centric evaluation. Because the phase-space entropy captures correlations between conjugate representations, it may extend to diagnosing dataset efficiency, generalization limits, and model vulnerabilities in broader AI systems. We envisage this framework as part of a more general, physics-informed language of data quality—one that quantifies how information, once measured, remains learnable.

**Supplementary Information.** Long-form derivations (with distributional conventions), ablations, and OTA hardware details are provided in the SI.

## Declarations

**Competing interests**: All authors declare no competing interests.
**Data availability**: The data used are all open source and can be found publicly.
**Code availability**: Husimi-based entropy auditing and mask generators can be found at https://doi.org/10.57760/sciencedb.31119 and https://doi.org/10.57760/sciencedb.31085.
**Author contribution**: conceptualization: Jun-Jie Zhang, Deyu Meng; theory: Jun-Jie Zhang, Long-Gang Pang; experiments: Xiu-Cheng Wang, Jun-Jie Zhang, Nan Cheng; writing: Jun-Jie Zhang, Xiu-Cheng Wang, Taijiao Du; supervision: Deyu Meng.

## References


[1] Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning, 1st edn. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA (2016). http://www.deeplearningbook.org

[2] LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015) https://doi.org/10.1038/nature14539

[3] Ng, A.: A chat with andrew on MLOps: From model-centric to data-centric AI (2021). Accessed: 2024-10-20

[4] Zha, H., He, Y., Liu, G., Li, J., Wang, D., Wang, H., Chen, Z., Liu, X., You, Y., Hsieh, C.-J., Zhang, C.: Data-centric AI: A new paradigm. arXiv preprint arXiv:2303.10158 (2023)

[5] Lustig, M., Donoho, D.L., Pauly, J.M.: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine **58**(6), 1182–1195 (2007) https://doi.org/10.1002/mrm.21391

[6] Kak, A.C., Slaney, M.: Principles of Computerized Tomographic Imaging. SIAM, Philadelphia (2001). https://doi.org/10.1137/1.9780898719277

[7] Punjani, A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A.: cryosparc: Algorithms for rapid unsupervised cryo-em structure determination. Nature Methods **14**, 290–296 (2017) https://doi.org/10.1038/nmeth.4169

[8] Marzetta, T.L.: Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Transactions on Wireless Communications **9**(11), 3590–3600 (2010) https://doi.org/10.1109/TWC.2010.092810.091092

[9] Jose, J., Ashikhmin, A., Marzetta, T.L., Vishwanath, S.: Pilot contamination and precoding in multi-cell TDD systems. IEEE Transactions on Wireless Communications **10**(8), 2640–2651 (2011) https://doi.org/10.1109/TWC.2011.060711.101155

[10] Larsson, E.G., Edfors, O., Tufvesson, F., Marzetta, T.L.: Massive MIMO for next generation wireless systems. IEEE Communications Magazine **52**(2), 186–195 (2014) https://doi.org/10.1109/MCOM.2014.6736761

[11] Alkhateeb, A., Leus, G., Heath, R.W.: Channel estimation and hybrid precoding for millimeter wave cellular systems. IEEE Journal of Selected Topics in Signal Processing **8**(5), 831–846 (2014) https://doi.org/10.1109/JSTSP.2014.2334278

[12] Gao, Z., Dai, L., Wang, Z., Chen, S., Wang, Z.: Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO. IEEE Journal on Selected Areas in Communications **34**(4), 998–1012 (2016) https://doi.org/10.1109/JSAC.2016.2545418

[13] He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.: Masked autoencoders are scalable vision learners. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 16000–16009 (2022)

[14] Bao, H., Dong, L., Piao, S., Wei, F.: Beit: BERT pre-training of image transformers. In: International Conference on Learning Representations (ICLR) (2022). arXiv:2106.08254

[15] Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., Efros, A.A.: Context encoders: Feature learning by inpainting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2536–2544 (2016). https://doi.org/10.1109/CVPR.2016.278

[16] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An image is worth 16x16 words: Transformers for image recognition at scale. In: International Conference on Learning Representations (ICLR) (2021)

[17] Xie, Z., Zhang, Z., Cao, Y., Lin, Y., Bao, J., Yao, Z., Dai, Q., Hu, H.: Simmim: A simple framework for masked image modeling. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9653–9663 (2022)

[18] Donoho, D.L.: Compressed sensing. IEEE Transactions on Information Theory **52**(4), 1289–1306 (2006) https://doi.org/10.1109/TIT.2006.871582

[19] Candès, E.J., Wakin, M.B.: An introduction to compressive sampling. IEEE Signal Processing Magazine **25**(2), 21–30 (2008) https://doi.org/10.1109/MSP.2007.914731

[20] Candès, E.J., Romberg, J.K., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Communications on Pure and Applied Mathematics **59**(8), 1207–1223 (2006) https://doi.org/10.1002/cpa.20124

[21] Candès, E.J., Tao, T.: Near-optimal signal recovery from random projections: Universal encoding strategies? IEEE Transactions on Information Theory **52**(12), 5406–5425 (2006) https://doi.org/10.1109/TIT.2006.885507

[22] Mallat, S.: A Wavelet Tour of Signal Processing: The Sparse Way, 3rd edn. Academic Press, Burlington, MA (2009)

[23] Recht, B., Fazel, M., Parrilo, P.A.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM Review **52**(3), 471–501 (2010) https://doi.org/10.1137/070697835

[24] Krause, A., Guestrin, C.: Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. In: Proceedings of the 22nd International Conference on Machine Learning (ICML) (2005)

[25] Krause, A., Singh, A., Guestrin, C.: Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. In: Journal of Machine Learning Research, vol. 9, pp. 235–284 (2008)

[26] Pukelsheim, F.: Optimal Design of Experiments. Classics in Applied Mathematics, vol. 50. SIAM, Philadelphia, PA (2006). https://doi.org/10.1137/1.9780898719100

[27] Atkinson, A.C., Donev, A.N., Tobias, R.D.: Optimum Experimental Designs, with SAS. Oxford Statistical Science Series, vol. 34, p. 528. Oxford University Press, Oxford; New York (2007). Includes SAS program code and companion website

[28] Cover, T.M., Thomas, J.A.: Elements of Information Theory, 2nd edn. Wiley Series in Telecommunications and Signal Processing. John Wiley & Sons, Hoboken, NJ (2006). https://doi.org/10.1002/047174882X . First published 1991; Second edition includes new chapters on network information theory and rate-distortion theory. https://doi.org/10.1002/047174882X

[29] Koh, P.W., Liang, P.: Understanding black-box predictions via influence functions. In: Proceedings of the 34th International Conference on Machine Learning (ICML), pp. 1885–1894 (2017)

[30] Ghorbani, A., Zou, J.: Data shapley: Towards data valuation using shapley value. In: Advances in Neural Information Processing Systems (NeurIPS) (2019). Also in ICML 2019 (PMLR 97: Data Shapley: Equitable Valuation of Data for Machine Learning)

[31] Mirzasoleiman, B., Bilmes, J., Leskovec, J.: Coresets for data-efficient training of machine learning models. In: Proceedings of the 37th International Conference on Machine Learning (ICML), pp. 6950–6960 (2020)

[32] Killamsetty, K., Sivasubramanian, D., Ramakrishnan, G., De, A., Iyer, R.: Glister: Generalization based data subset selection for efficient and robust learning. In: Proceedings of the AAAI Conference on Artificial Intelligence (AAAI) (2021)

[33] Killamsetty, K., Sivasubramanian, D., Ramakrishnan, G., De, A., Iyer, R.: Grad-match: Gradient matching based data subset selection for efficient deep model training. In: Proceedings of the 38th International Conference on Machine Learning (ICML) (2021)

[34] Nguyen, C.V., Hassner, T., Seeger, M., Archambeau, C.: LEEP: A new measure to evaluate transferability of learned representations. In: Proceedings of the 37th International Conference on Machine Learning (ICML), pp. 7294–7305 (2020)

[35] You, K., Liu, Y., Wang, J., Long, M.: LogME: Practical assessment of pre-trained models for transfer learning. In: Proceedings of the 38th International Conference on Machine Learning (ICML) (2021)

[36] Wigner, E.P.: On the quantum correction for thermodynamic equilibrium. Physical Review **40**(5), 749–759 (1932) https://doi.org/10.1103/PhysRev.40.749

[37] Husimi, K.: Some formal properties of the density matrix. Proceedings of the Physico-Mathematical Society of Japan. 3rd Series **22**, 264–314 (1940)

[38] Cohen, L.: Time–frequency distributions—a review. Proceedings of the IEEE **77**(7), 941–981 (1989) https://doi.org/10.1109/5.30749

[39] Flandrin, P.: Time-Frequency/Time-Scale Analysis. Academic Press, San Diego (1999)

[40] Schleich, W.P.: Quantum Optics in Phase Space. Wiley-VCH, Weinheim, Germany (2011). https://doi.org/10.1002/9783527635004 . https://doi.org/10.1002/9783527635004

[41] Shannon, C.E.: A mathematical theory of communication. Bell System Technical Journal **27**(3), 379–423 (1948)

[42] Wehrl, A.: General properties of entropy. Reviews of Modern Physics **50**(2), 221–260 (1978)

[43] Zurek, W.H., Paz, J.P.: Decoherence, chaos, and the second law. Physical Review Letters **72**(16), 2508–2511 (1994)

[44] Cook, R.L.: Stochastic sampling in computer graphics. ACM Transactions on Graphics **5**(1), 51–72 (1986)

[45] Bridson, R.: Fast poisson disk sampling in arbitrary dimensions. In: ACM SIGGRAPH 2007 Sketches (2007)

[46] Balanis, C.A.: Antenna Theory: Analysis and Design, 4th edn. Wiley, Hoboken, NJ (2016). https://doi.org/10.1002/9781118642061 . https://doi.org/10.1002/9781118642061

[47] Vinyals, O., Blundell, C., Lillicrap, T., Kavukcuoglu, K., Wierstra, D.: Matching networks for one shot learning. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 29, pp. 3630–3638 (2016)

[48] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234–241 (2015). Springer. LNCS 9351, doi:10.1007/978-3-319-24574-4_28

[49] Zhang, J.-J., Zhang, D.-X., Chen, J.-N., Pang, L.-G., Meng, D.: On the uncertainty principle of neural networks. iScience **28**(4), 112197 (2025) https://doi.org/10.1016/j.isci.2025.112197

[50] Zhang, J.-J., Meng, D.: Quantum-inspired analysis of neural network vulnerabilities: The role of conjugate variables in system attacks. National Science Review **11**(9), 141 (2024) https://doi.org/10.1093/nsr/nwae141

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SI.pdf