DEPTH CALIBRATES THE EFFECTIVE BANDWIDTH OF SINUSOIDAL NEURAL FIELDS

Anonymous authors
Paper under double-blind review

ABSTRACT

The frequency-scale hyperparameter of a sinusoidal neural field, ω_0 or $f_{\max}$, is commonly treated as the main control on what the network represents. The spectrum a trained network realizes departs from that nominal scale through weight rescaling and through the harmonics that layer composition generates, and how the two relate as the network deepens has not been characterized. We measure the realized spectrum directly, through an effective bandwidth B_{eff} defined as the 95% spectral-energy radius of the trained output on a dense grid, over roughly 7,000 fits across two datasets, two architectures and four depths. Near-optimal nominal values span 4–$10.7\times$ across depths while the realized B_{eff} spans only 1.5–$2.3\times$. Depth amplifies spectral expansion. At a fixed knob the realized bandwidth grows from 18 to 727, and the operational expansion boundary shifts down with depth, $[32, 16, 12, 12]$ on both datasets under a fixed $2\times$ criterion. A refined sweep resolves that boundary into a steep, localized crossover, and near-optimal nominal scales lie no farther than its near edge. The geometry supports a target-spectrum-conditioned selector. It uses one ratio, fit on two shallow depths of one dataset, with the signal bandwidth B_{sig} estimated from training coordinates only. Over the completed candidate grid its mean regret is 0.20 dB on held-out depths, 0.59 dB on a held-out dataset without fitting a target-side bandwidth map, 0.52 dB transferring to an architecture-specific SIREN map and 0.12 dB on a 32-image held-out cohort. Its worst case is 2.35 dB, at one frequency just past the expansion boundary. Across depth the effective bandwidth is the more stable quantity, and it is the one to calibrate. All experiments concern controlled 2D image regression with sinusoidal-activation networks.

1 INTRODUCTION

Implicit neural representations parameterize continuous signals with coordinate-to-value networks and have become a common representation for shapes, images and scenes (Park et al., 2019; Mescheder et al., 2019; Sitzmann et al., 2019; Mildenhall et al., 2020; Chen et al., 2021; Xie et al., 2022). For sinusoidal neural fields, spectral content is commonly controlled through one nominal frequency-scale hyperparameter. SIREN uses ω_0 to set the scale of its sinusoidal activations at initialization (Sitzmann et al., 2020), while frequency-bank architectures expose $f_{\max}$ as the upper limit of an explicit frequency set. Such frequency parameters are inherited by downstream systems for tasks including generative scene modeling and image compression (Chan et al., 2021; Mehta et al., 2021; Dupont et al., 2021; Strümpler et al., 2022). Both analysis and tuning practice revolve around this nominal initialization value.

Yet the nominal value is separated from the spectrum the trained network actually realizes by at least three mechanisms. The parameters can drift during training, learned dense-layer weights rescale what each frequency sees, and composition across layers generates harmonics far above the initialization band (Yüce et al., 2022; Novello et al., 2025). The same nominal scale can therefore correspond to substantially different realized spectra, particularly as depth changes. Existing theory characterizes how sinusoidal composition generates and can bound frequencies, but it does not identify, for standard unconstrained trained networks, the depth-dependent realized-bandwidth regime associated with off-grid generalization.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

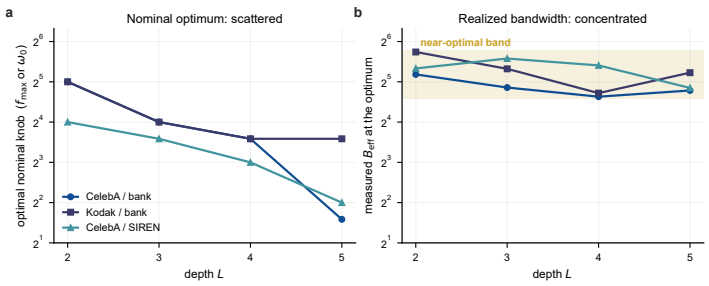


Figure 1: The nominal knob versus the realized spectrum. (a) The generalization-optimal nominal value scatters across depths, architectures, and datasets. (b) The measured B_{eff} of the same optimal configurations concentrates in a narrow band per setting.

We address this gap by measuring the realized spectrum directly. We define the effective bandwidth B_{eff} of a trained network as the radius containing 95% of the spectral energy of its output on a dense evaluation grid, and the signal bandwidth B_{sig} as the same measurement on the target, estimated from training coordinates only. With this instrument, we run roughly 7,000 fits across two datasets, two architectures, and four depths, together with a frozen-frequency control arm that pins down what the measurement means (Section 3). The measurements give a depth-dependent geometry linking the nominal scale to the realized bandwidth, and three findings follow from it (Figure 1).

First, the realized spectrum organizes generalization where the nominal knob does not (Section 4). The generalization-optimal nominal value moves with depth, from $f^*_{\max} = 32$ at $L = 2$ down to 3 at $L = 5$ on CelebA, and its deep-end value is unstable across images. Indexed by measured B_{eff} instead, the near-optimal models of all depths concentrate: within each evaluated setting, near-optimal nominal values span 4–10.7$\times$ while realized B_{eff} spans only 1.5–2.3$\times$.

Second, depth amplifies spectral expansion (Section 5). At a fixed knob, the realized bandwidth grows with depth, from 18 to 727 on CelebA at $f_{\max} = 16$, and the expansion boundary along the frequency axis shifts to lower nominal values as depth grows, $[32, 16, 12, 12]$ on both datasets under a fixed $2\times$ criterion. A refined frequency sweep resolves this boundary into a steep, localized crossover, and near-optimal nominal scales lie no farther than the near edge of rapid expansion. Layerwise measurements support a cascade-like amplification account, with flat per-layer bandwidth below the boundary and late-layer amplification past it.

Third, the measurements support an empirical, target-conditioned recalibration rule (Section 6). A single ratio between realized and signal bandwidth, fit on two shallow depths of one dataset, achieves low mean regret, the test-PSNR gap to the best grid candidate, across held-out depths and transfer settings, including an architecture-specific SIREN bandwidth map and three confirmation runs executed after every analysis definition was fixed, while exhibiting one identifiable failure at a spectral-expansion shoulder. For the frequency-bank architecture, deepening without recalibrating the frequency setting can move the model across the measured expansion boundary.

Our contributions are:

- a systematic, depth-resolved measurement of realized bandwidth across nominal frequency, depth, architecture and dataset, showing that within each evaluated setting near-optimal models occupy a narrower range in B_{eff} than in nominal frequency, 1.5–2.3$\times$ against 4–10.7$\times$.
- a depth-dependent spectral-expansion regime whose onset shifts toward lower nominal frequencies as depth grows, $[32, 16, 12, 12]$ under a fixed $2\times$ criterion, with near-optimal nominal scales no farther than its near edge and replicated layerwise evidence for late-layer amplification.
- a target-spectrum-conditioned selection rule, built from one calibration ratio and a signal bandwidth estimated from training coordinates only, evaluated on held-out depths and a held-out dataset, and transferring to an architecture-specific SIREN bandwidth map.

The scope of every result, including the fixed image subsets and the sinusoidal-family boundary, is made explicit in Section 7.

2 Related work

Sinusoidal fields and their frequency setting. SIREN (Sitzmann et al., 2020) established sinusoidal activations with a frequency scale ω_0 and a matched initialization. Random Fourier features (Tancik et al., 2020) place the frequencies in the input encoding instead, and analyze the scale through the neural tangent kernel (Jacot et al., 2018; Arora et al., 2019), and the same scale is tuned in multi-scale PDE solvers (Wang et al., 2021). In both lines of work, and in the tuning practice built on them, the frequency setting is the nominal initialization value. Our work is complementary: we measure the spectrum the trained network realizes and show that the nominal setting must be recalibrated across depth.

Frequency generation by layer composition. Yüce et al. (2022) showed that the functions expressible by a sinusoidal INR are built from integer harmonic combinations of the initial frequencies, with a frequency support that grows with depth. TUNER (Novello et al., 2025) gives an amplitude-phase expansion for sinusoidal MLPs and uses it to bound the spectrum during training. A Fourier-series view reaches a compatible single-layer picture (Benbarka et al., 2022). These works supply the mechanism-level theory for why composition expands the spectrum. What they do not supply, and what we contribute, is the measured object itself: the depth-resolved realized-bandwidth regime of near-optimal configurations, the crossover and boundary maps, and a calibration ratio that transfers across the evaluated settings. TUNER constrains the spectrum during training, while our rule selects a nominal setting from measurements. The two are potentially complementary, a combination we do not evaluate.

Bandlimited by construction. BACON (Lindell et al., 2022) designs coordinate networks whose outputs are bandlimited by construction. Multiplicative filter networks constrain the compositional structure itself (Fathony et al., 2021; Shekarforoush et al., 2022), and adaptive or multiresolution encodings address fine detail (Martel et al., 2021; Müller et al., 2022). These methods provide design-side answers to spectral control. Our answer is measurement-side for unconstrained sinusoidal architectures, quantifying what standard architectures realize and selecting settings accordingly.

Spectral bias and its remedies. The spectral-bias literature established that neural networks learn low frequencies first (Rahaman et al., 2019; Xu et al., 2020; Basri et al., 2019; 2020) and, for coordinate networks, prescribed added bandwidth as the fix (Tancik et al., 2020). Progressive encodings instead anneal or unmask positional-encoding frequency bands during optimization (Park et al., 2021; Lin et al., 2021; Hertz et al., 2021). That prescription is tuned for fitting. Our measurements locate its cost for off-grid generalization, a depth-dependent boundary past which added nominal bandwidth becomes realized spectrum far beyond the signal. Activation families beyond sinusoids, including wavelet, Gaussian, and variable-periodic designs (Saragadam et al., 2023; Ramasinghe & Lucey, 2022; Liu et al., 2024), are outside our measured scope and are discussed as open questions in Section 7.

3 Setup: tasks, architectures, and the measured spectrum

3.1 Tasks and data

We study coordinate regression on images. A network $f_\theta : [-1, 1]^2 \to \mathbb{R}^3$ receives a pixel coordinate and returns its RGB value. Each image is resized to 128×128. We train on a random 25% of the pixel coordinates and evaluate PSNR on the held-out 75%, with predictions clipped to $[0, 1]$. The same fixed coordinate permutation (seed 123) is used for every image and configuration. Reported seeds vary the network initialization, not the train/test split. This split measures off-grid generalization directly: three quarters of the image is never seen during training.

We use two datasets: CelebA (Liu et al., 2015) (8 images, 3 seeds per configuration) and Kodak (8 images, 2 seeds). In each dataset the eight images are simply the first eight in filename order. This subset was fixed before any sweep was run (Section 7). In total the study covers roughly 7,000

instrumented fits, including confirmation and control runs executed after the analysis definitions of Section 3.4 were frozen (Appendix H). We call these frozen-protocol runs. Aggregate numbers are means over images and seeds.

3.2 Two sinusoidal architectures

We evaluate two representatives of the sinusoidal family, neither introduced as a contribution of this paper. Both expose one nominal frequency hyperparameter, the knob whose meaning we measure.

Frequency-bank network. Each activation unit computes $\sigma(x) = \sum_{i=1}^{8} b_i \sin(\omega_i x + \phi_i)$. In the main arm (Section 3.5), the eight frequencies are fixed to $\mathrm{geomspace}(2, f_{\max}, 8)$, while amplitudes, phases, and dense-layer parameters remain trainable. The top frequency $f_{\max}$ is the nominal knob, and every bank shares the $\omega = 2$ endpoint. The network stacks L blocks of a dense layer (width 256) followed by this activation, with a linear head.

SIREN. The standard architecture of Sitzmann et al. (2020): L sinusoidal hidden layers of width 256 computing $\sin(\omega_0(Wx + b))$, followed by a linear head, with all parameters trainable and fixed ω_0 as the nominal knob. Initialization follows the original scheme (Appendix G).

All runs use Adam (Kingma & Ba, 2015) with the learning rate exponentially decayed from 10^{-3} to approximately 5×10^{-5} over 800 iterations, at depths $L \in \{2, 3, 4, 5\}$. All reported PSNR and spectral measurements use the model after the last of the 800 iterations. We never look at the held-out pixels during training, neither to stop early nor to pick the best checkpoint. They enter only at evaluation time.

3.3 Effective bandwidth and signal bandwidth

The nominal knob sets only the initialization. What the trained network realizes spectrally is a separate quantity, which we measure directly. After training, we sample the network on a dense 512^2 grid over $[-1, 1]^2$ and compute the channel-summed 2D power spectrum $P(\mathbf{k})$ with the DC component removed. The *effective bandwidth* B_{eff} is the radial frequency below which 95% of the spectral energy lies,

$$B_{\text{eff}} \;=\; \min\Big\{\, r \;:\; \textstyle\sum_{\|\mathbf{k}\| \le r} P(\mathbf{k}) \;\ge\; 0.95 \sum_{\mathbf{k}} P(\mathbf{k}) \,\Big\}. \tag{1}$$

B_{eff} is measured in the frequency units of the evaluation grid. It is not on the same scale as $f_{\max}$ or ω_0. The knob acts on the network through trained weights, while B_{eff} is measured on the output, and the two numbers should not be compared directly. We treat B_{eff} as an instrument. It is a fixed and reproducible readout of what a trained network produces, and it does not estimate any theoretical quantity. The spectrum is computed from the raw network output, without clipping to $[0, 1]$. Clipping changes B_{eff} by at most about 7% (Appendix E). Appendix D shows how the selector behaves when the 95% threshold is moved.

The *signal bandwidth* B_{sig} applies the same measurement to the target image. Computing it on the full image would touch the held-out pixels. We therefore estimate B_{sig} from the training coordinates only: the observed 25% of pixels are linearly interpolated on the native 128^2 grid, with nearest-neighbor filling outside the convex hull, and Equation 1 is applied to the result. The interpolated image is smoother than the true one, and this estimator therefore measures low. Calibrating the selector's ratio (Section 6.1) with the same estimator compensates for this bias in the evaluated settings. The estimate is not fragile. In the four protocols checked in Appendix D, nearest-neighbor filling of the missing pixels leads to the same selections as linear interpolation, and re-drawing the observed 25% of pixels five different ways also leaves every selection unchanged. The full-image value serves only as an oracle diagnostic in the appendix, never as part of any selection.

3.4 Fixed analysis definitions and completed candidate grids

Near-optimal set. For each depth, the set of nominal values whose mean test PSNR lies within 0.5 dB of the best evaluated value at that depth. A 1.0 dB variant appears in Appendix D.

Expansion boundary. The last candidate before the first adjacent pair whose B_{eff} ratio reaches $2\times$. The boundary is a property of the evaluated grid (Section 5).

Candidate grid. The bank sweep evaluates $f_{\max} \in \{3, 4, 6, 8, 12, 16, 20, 24, 32, 48, 64\}$ at every depth, and the SIREN sweep evaluates $\omega_0 \in \{4, 8, 12, 16, 20, 24, 32, 48, 64, 128\}$. The bank value $f_{\max} = 20$ was added after the first sweeps to resolve the $L = 3$ transition, and we call the resulting grid the completed grid. Selection results in Section 6 are defined on these grids and only on them.

3.5 Frozen-frequency control

Because the nominal frequencies are trainable parameters in the bank architecture, B_{eff} could in principle reflect a moving frequency set and not what training does with a fixed one. We control for this directly. In the dense CelebA $L = 3$ control sweep, every evaluated frequency is run in two arms. In the frozen arm the activation frequencies stay at their initial values. In the free arm they are trained like every other parameter. The two arms agree. Across the evaluated frequencies, the aggregate test PSNR of the two arms differs by at most 0.30 dB. In the free arm the frequencies move by only 0.34% of their initial values on average, with a per-run 95th percentile of 0.82% and a maximum of 1.12% (Appendix A). The bandwidth patterns of the main arm therefore cannot come from drift of the nominal frequencies. The frozen arm is the main arm throughout the paper.

4 The realized spectrum organizes generalization

4.1 The nominal optimum scatters across depth

Sweeping the nominal knob at a fixed depth reproduces the known picture. On CelebA at $L = 3$, train PSNR rises monotonically with $f_{\max}$, up to 42.9 dB at $f_{\max} = 48$, while test PSNR peaks at 21.6 dB at $f_{\max} = 16$ and then collapses to 7.8 dB, the inverted U already documented in prior work (Tancik et al., 2020; Yüce et al., 2022).

The difficulty appears when depth varies. The nominal optimum moves: $f^*_{\max} = 32/16/12/3$ on CelebA and $32/16/12/12$ on Kodak for $L = 2/3/4/5$, and $\omega^*_0 = 16/12/8/4$ for SIREN (Table 3, Figure 2a). Within each evaluated setting the near-optimal nominal values span 4–$10.7\times$ across depths. The deep end is also unstable at the level of single images. Per-image optima at $L = 4$ and $L = 5$ scatter across several nominal values, and an image-level bootstrap reproduces the deep-end argmax less often than the shallow one (Appendix D). A practitioner who tunes the knob at one depth therefore learns little about the next depth.

4.2 The realized bandwidth is far more stable

The picture changes when the same models are indexed by what they realize, not by their initialization setting. At each depth we take the near-optimal set, all nominal values within 0.5 dB of the best evaluated value, and take the measured B_{eff} of those models (Figure 2b). Within each evaluated setting, realized bandwidth varies substantially less across near-optimal depths than the nominal frequency scale: the near-optimal nominal values span $10.7\times$/$10.7\times$/$4.0\times$ (CelebA/Kodak/SIREN), while the corresponding B_{eff} spans only $1.52\times$/$2.32\times$/$2.15\times$. Three frozen-protocol confirmation grids, SIREN on Kodak and held-out image cohorts of both datasets, show the same pattern (Appendix H).

B_{eff} is therefore a more stable organizing variable across depth, while the nominal knob is an architecture- and depth-dependent proxy for it. The bands do differ between settings (Table 3), and they are not a universal constant. The result concerns dispersion within each setting, and it holds in all four evaluated settings. The spans describe how far the numbers spread on the grids we evaluated. They do not show that a particular B_{eff} causes good generalization, only that the good configurations share a narrow B_{eff} band.

The deep-end instability in Section 4.1 has a simple interpretation in the measured space. At $L = 4$ and $L = 5$ on CelebA, nominal values from 3 to 12 all realize B_{eff} between roughly 22 and 36. Several distinct knob settings map to nearly the same realized spectrum, and in this sense the parameterization is many-to-one. We refer to this plateau of equivalent settings as the deep-end degeneracy. Small sampling or optimization perturbations can therefore change the nominal argmax while leaving the realized bandwidth band largely unchanged. The bootstrap agrees with this interpretation. The two depths with wide plateaus, $L = 4$ and $L = 5$, are also the two depths

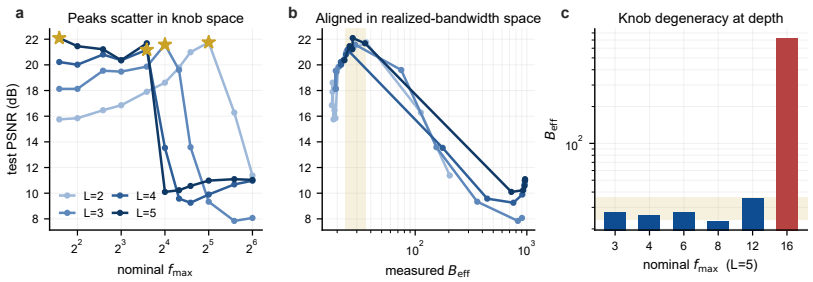


Figure 2: The realized spectrum organizes generalization. (a) Test PSNR against the nominal knob forms a fan of curves whose peaks shift with depth. (b) Indexed by measured B_{eff}, the near-optimal models of all depths concentrate in a narrow band. (c) At deep L, several nominal values map to nearly the same B_{eff}, consistent with the observed deep-end scatter of the nominal argmax.

where the argmax was least stable under resampling, at 82% and 70% against 92–100% at shallow depths (Appendix D).

4.3 Input-encoding control

A random-Fourier-feature network (Tancik et al., 2020) places its frequencies in the input encoding, not in the activations. In a parameter-matched control (ReLU width solved per depth to match the bank's trainable parameter count within 0.4%), no bank-like depth-dependent shift appears. With σ the scale of the Gaussian encoding frequencies, the optimum is $\sigma^* = 1/1/1/0.5$ against the $10.7\times$ shift of the bank optimum, and $\sigma = 1$ remains within the 0.5 dB near-optimal set at every depth (Appendix F).

5 Depth-amplified spectral expansion and its operational crossover

Why does the nominal optimum move? The measurement gives a direct answer. At a fixed knob, depth expands what the network realizes. Holding $f_{\max} = 16$, the output B_{eff} grows from 18 to 29, 176, and 727 for $L = 2, 3, 4, 5$ on CelebA, and from 21 to 40, 202, and 782 on Kodak (Figure 3a,b). The shallow steps are mild. The deep steps are not: between $L = 3$ and $L = 5$ the realized bandwidth grows by about $25\times$ on CelebA and $20\times$ on Kodak, approaching the Nyquist limit of the evaluation grid. A deeper network with the same knob is spectrally a different network, which is exactly what a depth-independent tuning practice ignores.

5.1 The expansion boundary shifts down with depth

Along the frequency axis, the expansion is strongly nonuniform for the bank. B_{eff} tracks the knob closely up to a candidate value and then jumps. On CelebA at $L = 3$ the sequence is 29 at $f_{\max} = 16$, 74.9 at 20, and 155 at 24: past the boundary, one grid step multiplies the realized bandwidth. We call the first candidates past the boundary, where B_{eff} has begun to jump but has not yet saturated, the expansion shoulder. Under the fixed $2\times$ rule of Section 3.4, both datasets yield $[32, 16, 12, 12]$ for $L = 2/3/4/5$. Under the $3\times$ criterion, the $L = 3$ boundary disappears in both datasets after adding $f = 20$. The exact boundary location thus depends on the jump threshold and grid resolution, while the downward shift is clear under the fixed $2\times$ criterion.

Two facts connect the boundary to generalization. First, on the completed grid, every near-optimal configuration in Table 3 lies on the pre-expansion side of its boundary. Second, at greater depths several nominal settings occupy the same near-optimal plateau below the boundary, which is the degeneracy of Section 4.2. The two held-out image cohorts reproduce the same boundary under the same criterion (Appendix H).

A refined sweep resolves what the boundary rule detects. Adding $f_{\max} \in \{14, 18, 22\}$ at $L = 3$ on both datasets yields a seven-point local curve on which no adjacent B_{eff} ratio reaches $2\times$,

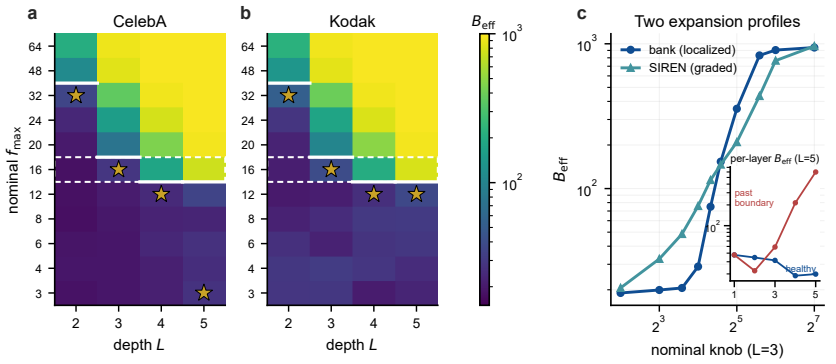


Figure 3: Depth-amplified spectral expansion. (a,b) B_{eff} as a function of depth and nominal frequency on CelebA and Kodak. The expansion boundary under the $2\times$ rule (marked) shifts to lower nominal values with depth, and on this grid near-optimal configurations lie on its pre-expansion side. At fixed $f_{\max} = 16$ (outlined row), the measured B_{eff} rises from 18 to 727 on CelebA as depth increases. (c) Bank versus SIREN growth profiles: a localized steep crossover for the bank, a steep but graded expansion for SIREN. Inset: per-layer bandwidth for a pre-boundary and a past-boundary configuration (single image and seed, illustrative; replication across images and seeds in Appendix H).

because the inserted candidates split the former coarse-grid jump. The grid-normalized slope $s_i = \Delta \log B_{\text{eff}} / \Delta \log f_{\max}$ rises from 0.7–1.3 below the shoulder to a peak of 5.66 on CelebA and 4.25 on Kodak, both on the $[18, 20]$ pair (Appendix H). The expansion is therefore a steep crossover at the same location on both datasets, and the boundary in Section 3.4 is its operational, grid-level diagnostic. The refined aggregate optimum moves to $f_{\max} = 18$ on CelebA, better than 16 by 0.40 dB, and stays at 16 on Kodak. The refined near-optimal set $\{16, 18\}$ ends at the near endpoint of the peak-slope interval $[18, 20]$.

5.2 SIREN EXPANDS STEEPLY BUT GRADUALLY

The same depth trend holds for SIREN, although the transition is graded, not localized. The $2\times$ rule fires only at scattered positions ($\omega_0 = 64/32/32/4$ across depths), the $3\times$ rule fires nowhere, and the largest adjacent-step ratio stays between 2.0 and 3.0 on both the CelebA and the Kodak-SIREN confirmation grids, against ratios of over $20\times$ for the bank (Appendices D and H). The sharper bank transition may be related to its discrete geometric frequency set, although the present evidence does not identify the mechanism.

5.3 LAYERWISE MEASUREMENTS SUPPORT A CASCADE-LIKE ACCOUNT

Measuring the spectrum layer by layer shows where the expansion occurs inside the network (Figure 3c, inset). In pre-boundary configurations the per-layer bandwidth stays flat at roughly 20–40 through the whole stack. Past the boundary the expansion concentrates in the late layers and amplifies stage by stage: at $L = 5$ with $f_{\max} = 16$ the per-layer sequence is $38 \to 22 \to 49 \to 214 \to 599$. The first layer changes little across knob settings, which is consistent with dense-weight rescaling. These layerwise measurements support a cascade-like amplification account, in which depth supplies more stages for harmonics to compound (Yüce et al., 2022; Novello et al., 2025). The parameter-matched Fourier-feature control of Section 4.3 shows no such depth shift, which is consistent with composition inside the network but does not isolate it. A repeated run, four images and two seeds per configuration on both datasets, counts the pattern: a final layer with at least twice the first layer's B_{eff} appears in 32 of 32 past-boundary image-seed pairs and 0 of 32 pre-boundary pairs (Appendix H).

Table 1: Selector regret on the seven held-out protocols: mean point estimate with 95% image-level bootstrap CI, and per-depth maximum where the text quotes it. The last three are frozen-ρ confirmations (Appendix H). Per-depth values appear as points in Figure 4b.

Protocol	Held out from calibration	Mean (dB)	95% CI	Max (dB)
CelebA, $L \in \{4, 5\}$	depths	0.20	$[0.00, 0.65]$	–
Kodak, own map	dataset	0.49	$[0.02, 1.31]$	–
Kodak, source map	dataset, no target-side map	0.59	$[0.02, 1.76]$	2.35
SIREN	architecture	0.52	$[0.32, 0.71]$	–
Kodak-SIREN, frozen ρ	dataset and architecture	0.06	$[0.00, 0.37]$	0.16
Kodak 9–24 cohort, frozen ρ	fresh images	0.08	$[0.00, 0.39]$	0.31
CelebA 9–40 cohort, frozen ρ	fresh images, source dataset	0.12	$[0.00, 0.45]$	0.48

6 A TARGET-SPECTRUM-CONDITIONED EMPIRICAL SELECTOR

A single calibrated ratio converts the measurements into a selection rule.

6.1 RULE AND PROTOCOL

The measurements of Sections 4 and 5 suggest a simple recipe: aim the realized bandwidth at a fixed fraction of the signal bandwidth. We fit the ratio $\rho = B^*_{\text{eff}}/B_{\text{sig}}$ on CelebA at the two shallow depths only ($L \in \{2, 3\}$), which gives $\rho = 0.78$ with the train-only B_{sig} estimator of Section 3.3. A bandwidth map is the table of measured B_{eff} for every candidate at one depth, as in Tables 4 and 5. The selector then picks, among the candidates of such a map, the one whose B_{eff} is closest to $\rho\, B_{\text{sig}}$ of the target image set. Everything downstream of this fit is held out: the deeper CelebA depths, the entire Kodak dataset, and the SIREN architecture never touch the calibration.

Evaluation is selection-level. On the completed candidate grid of Section 3.4, the regret of a selection is the mean test PSNR gap to the best evaluated candidate at that depth. We report point estimates with 95% image-level bootstrap intervals (2,000 resamples). Appendix D lists what the bootstrap recomputes in each protocol.

The protocols differ in what is trained on the target. In the own-map protocols, each candidate is trained on the target images and its B_{eff} is measured there. This gives the target its own bandwidth map. Selection then uses that map, the train-only B_{sig}, and the frozen ρ. It never looks at test PSNR on the target, and it never re-fits ρ. The source-map protocol reuses the CelebA map in place of the target's own map. This avoids the target-side candidate fits needed to build a map, while the selected network is still trained on the target images. The reuse pays at the shoulder (Section 6.3), in exchange for an eleven-fold saving in target-side fits (Appendix D).

6.2 HELD-OUT RESULTS

Four original protocols and three frozen-protocol confirmations probe complementary held-out axes. Table 1 reports the seven mean regrets with bootstrap intervals, and Figure 4b overlays the per-depth values. The source-map maximum of 2.35 dB is the shoulder failure analyzed in Section 6.3. The two image cohorts show that the ratio holds on fresh images of both datasets (Appendix H).

An absolute rule is the natural baseline. Fixing the target bandwidth to the CelebA-optimal value ($B = 32.7$, the mean B_{eff} at the two calibration optima) fails on Kodak at every depth (mean regret 2.69 dB, maximum 3.21 dB, own-map protocol), while matching the relative rule on the CelebA held-out depths. Relative scaling transfers better in the evaluated settings.

6.3 FAILURE ANALYSIS: THE EXPANSION SHOULDER

The single large miss is informative. At Kodak $L = 3$ the source-map selector picks $f_{\max} = 20$: on the CelebA map this candidate realizes $B_{\text{eff}} = 74.9$, closest to the Kodak target of $\rho\, B_{\text{sig}} \approx 64$ (Figure 4a). On Kodak itself the same nominal value realizes $B_{\text{eff}} = 101.8$, past the target, and the better choice is $f_{\max} = 16$. The miss costs 2.35 dB. The selection is unchanged across five B_{sig}-estimation masks, which rules out a mask artifact. Near the expansion shoulder of Section 5.1,

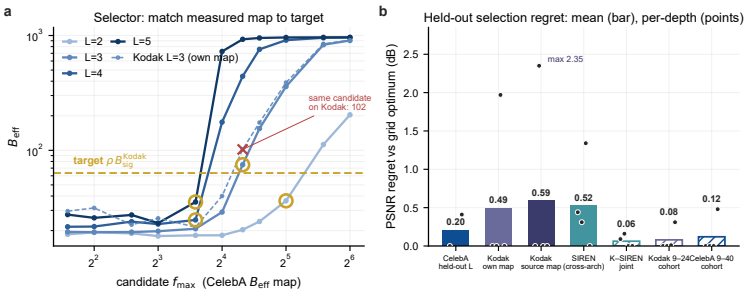


Figure 4: The empirical selector. (a) The rule picks the candidate whose measured B_{eff} on the source (CelebA) map is closest to the target $\rho\, B_{\text{sig}}$ (gold circles). At $L = 3$ it lands on the expansion shoulder, the one observed failure: on the Kodak $L = 3$ map (dashed) the same candidate realizes $B_{\text{eff}} = 102$ (cross). (b) Mean selection regret for seven held-out protocols (bars), with per-depth regrets overlaid as points. The three rightmost are frozen-ρ confirmations, and Kodak-SIREN uses a target-specific map without recalibration.

realized bandwidth changes fastest and differs most between settings. A source map is least transferable exactly there, and one should probe the target's own map near the shoulder. Appendix D reports the remaining robustness checks.

7 Discussion and Limitations

Spectral bias, seen from the generalization side. The spectral-bias literature prescribes added bandwidth as the remedy for slow high-frequency fitting (Rahaman et al., 2019; Sitzmann et al., 2020; Tancik et al., 2020). Our measurements quantify the other side of this prescription. Past a depth-dependent boundary, added nominal bandwidth turns into realized spectrum far beyond the signal, and off-grid reconstruction collapses while training accuracy still improves. The same property that slows fitting also acts as implicit regularization at unseen coordinates, and the boundary maps in Section 5 indicate where the protection ends.

Limitations. First, the selector picks from an evaluated grid. It does not predict the optimum between grid points, and the boundary it relies on is a grid-level diagnostic, not a grid-independent point (Section 5.1). Second, the source-map protocol misses at the expansion shoulder, at a cost of 2.35 dB at one depth. Third, the layerwise measurements support a cascade-like account but do not establish the mechanism. Fourth, at the 99% energy threshold the selector degrades (Appendix D). Fifth, the original grids use fixed eight-image subsets. The one-seed held-out cohorts of Appendix H reproduce the boundary and the low selector regret on sixteen Kodak and thirty-two CelebA images, but both datasets remain fixed choices. Finally, all experiments concern controlled 2D image regression at a single sampling regime (25% training coordinates) and are limited to sinusoidal-activation INRs. Whether the same expansion geometry appears in WIRE, Gaussian, or FINER-style activations (Saragadam et al., 2023; Ramasinghe & Lucey, 2022; Liu et al., 2024) is an open question, as is behavior on 3D and signed-distance tasks. Training difficulty at extreme configurations is a separate axis that we do not study. Together, the generalization and trainability constraints suggest a two-sided usable window for the sinusoidal family.

8 Conclusion

The frequency hyperparameter of a sinusoidal neural field is not the bandwidth it realizes. Across two datasets, two architectures and four depths, near-optimal nominal values span 4–$10.7\times$ while the realized near-optimal bandwidth stays within a 1.5–$2.3\times$ band per setting. At a fixed knob, depth amplifies the realized spectrum toward the grid Nyquist limit, and the expansion boundary shifts down with depth into a steep crossover that near-optimal nominal scales sit at or before. With one calibration ratio frozen, the selector achieves low mean regret on held-out depths, datasets, image cohorts and an architecture-specific SIREN map. Deepening such a network calls for recalibrating its frequency setting against the realized spectrum.

REPRODUCIBILITY STATEMENT

Section 3 gives the two datasets, the two architectures and the four depths in full, together with the definition of the effective bandwidth and the analysis settings that were fixed before the sweeps. The selection rule of Section 6 is stated with its single calibration ratio, the two shallow depths and the one dataset that ratio was fit on, and the held-out depths, dataset, architecture and image cohorts it is evaluated on. Appendix A through Appendix C give the completed candidate grids behind every number, and Appendix D lists the robustness checks. We will release the code, the result records behind every number in the paper and the scripts that produce every figure when the paper is published.

AI USAGE DISCLOSURE

Large language models were used throughout this project under the direction of the authors, in three ways. They took part in research ideation and execution, proposing hypotheses and writing the analysis and plotting code. They drafted and revised every section of the text. They also carried out an internal adversarial review of statements, numbers and experimental procedures, which included cross-checking the manuscript against the archived result records. The authors set the research questions, made the experimental design decisions, and verified every reported number against those records. The authors take full responsibility for the content of this paper.

REFERENCES

Sanjeev Arora, Simon S. Du, Wei Hu, Zhiyuan Li, and Ruosong Wang. Fine-grained analysis of optimization and generalization for overparameterized two-layer neural networks. In *International Conference on Machine Learning (ICML)*, pp. 322–332, 2019.

Ronen Basri, David Jacobs, Yoni Kasten, and Shira Kritchman. The convergence rate of neural networks for learned functions of different frequencies. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

Ronen Basri, Meirav Galun, Amnon Geifman, David Jacobs, Yoni Kasten, and Shira Kritchman. Frequency bias in neural networks for input of non-uniform density. In *International Conference on Machine Learning (ICML)*, pp. 685–694, 2020.

Nuri Benbarka, Timon Höfer, Hamd ul Moqeet Riaz, and Andreas Zell. Seeing implicit neural representations as Fourier series. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 2283–2292, 2022.

Eric R. Chan, Marco Monteiro, Petr Kellnhofer, Jiajun Wu, and Gordon Wetzstein. pi-GAN: Periodic implicit generative adversarial networks for 3D-aware image synthesis. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5795–5805, 2021.

Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning continuous image representation with local implicit image function. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8628–8638, 2021.

Emilien Dupont, Adam Goliński, Milad Alizadeh, Yee Whye Teh, and Arnaud Doucet. COIN: COmpression with implicit neural representations. *arXiv preprint arXiv:2103.03123*, 2021.

Rizal Fathony, Anit Kumar Sahu, Devin Willmott, and J. Zico Kolter. Multiplicative filter networks. In *International Conference on Learning Representations (ICLR)*, 2021.

Amir Hertz, Or Perel, Raja Giryes, Olga Sorkine-Hornung, and Daniel Cohen-Or. SAPE: Spatially-adaptive progressive encoding for neural optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 8820–8832, 2021.

Arthur Jacot, Franck Gabriel, and Clément Hongler. Neural tangent kernel: Convergence and generalization in neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

Chen-Hsuan Lin, Wei-Chiu Ma, Antonio Torralba, and Simon Lucey. BARF: Bundle-adjusting neural radiance fields. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5721–5731, 2021.

David B. Lindell, Dave Van Veen, Jeong Joon Park, and Gordon Wetzstein. BACON: Band-limited coordinate networks for multiscale scene representation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16231–16241, 2022.

Zhen Liu, Hao Zhu, Qi Zhang, Jingde Fu, Weibing Deng, Zhan Ma, Yanwen Guo, and Xun Cao. FINER: Flexible spectral-bias tuning in implicit NEural representation by variable-periodic activation functions. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2713–2722, 2024.

Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In *IEEE International Conference on Computer Vision (ICCV)*, pp. 3730–3738, 2015.

Julien N.P. Martel, David B. Lindell, Connor Z. Lin, Eric R. Chan, Marco Monteiro, and Gordon Wetzstein. ACORN: Adaptive coordinate networks for neural scene representation. *ACM Transactions on Graphics*, 40(4):58:1–58:13, 2021.

Ishit Mehta, Michaël Gharbi, Connelly Barnes, Eli Shechtman, Ravi Ramamoorthi, and Manmohan Chandraker. Modulated periodic activations for generalizable local functional representations. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 14194–14203, 2021.

Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3D reconstruction in function space. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4455–4465, 2019.

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *European Conference on Computer Vision (ECCV)*, 2020.

Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41(4):102:1–102:15, 2022.

Tiago Novello, Diana Aldana, Andre Araujo, and Luiz Velho. Tuning the frequencies: Robust training for sinusoidal neural networks. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3071–3080, 2025.

Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 165–174, 2019.

Keunhong Park, Utkarsh Sinha, Jonathan T. Barron, Sofien Bouaziz, Dan B. Goldman, Steven M. Seitz, and Ricardo Martin-Brualla. Nerfies: Deformable neural radiance fields. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5845–5854, 2021.

Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, and Aaron Courville. On the spectral bias of neural networks. In *International Conference on Machine Learning (ICML)*, pp. 5301–5310, 2019.

Sameera Ramasinghe and Simon Lucey. Beyond periodicity: Towards a unifying framework for activations in coordinate-MLPs. In *European Conference on Computer Vision (ECCV)*, pp. 142–158, 2022.

Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, and Richard G. Baraniuk. WIRE: Wavelet implicit neural representations. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 18507–18516, 2023.

Shayan Shekarforoush, David B. Lindell, David J. Fleet, and Marcus A. Brubaker. Residual multiplicative filter networks for multiscale reconstruction. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 8550–8563, 2022.

Vincent Sitzmann, Michael Zollhöfer, and Gordon Wetzstein. Scene representation networks: Continuous 3D-structure-aware neural scene representations. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

Vincent Sitzmann, Julien N.P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 7462–7473, 2020.

Yannick Strümpler, Janis Postels, Ren Yang, Luc Van Gool, and Federico Tombari. Implicit neural representations for image compression. In *European Conference on Computer Vision (ECCV)*, pp. 74–91, 2022.

Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 7537–7547, 2020.

Sifan Wang, Hanwen Wang, and Paris Perdikaris. On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, 384:113938, 2021.

Yiheng Xie, Towaki Takikawa, Shunsuke Saito, Or Litany, Shiqin Yan, Numair Khan, Federico Tombari, James Tompkin, Vincent Sitzmann, and Srinath Sridhar. Neural fields in visual computing and beyond. *Computer Graphics Forum*, 41(2):641–676, 2022.

Zhi-Qin John Xu, Yaoyu Zhang, Tao Luo, Yanyang Xiao, and Zheng Ma. Frequency principle: Fourier analysis sheds light on deep neural networks. *Communications in Computational Physics*, 28(5):1746–1767, 2020.

Gizem Yüce, Guillermo Ortiz-Jiménez, Beril Besbinar, and Pascal Frossard. A structured dictionary perspective on implicit neural representations. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 19206–19216, 2022.

A Frozen-frequency control: full numbers

Table 2 reports the dense CelebA $L = 3$ control sweep in both arms. The maximum aggregate test-PSNR gap between arms across the evaluated frequencies is 0.30 dB at the reported precision. In the free arm, the activation frequencies drift by 0.34% of their initial values on average, with a per-run 95th percentile of 0.82% and a maximum of 1.12%. Relative to the nominal top frequency $f_{\max}$, the mean drift is 0.12%. Under either normalization, the frequency set is close to stationary, which supports interpreting the main-arm B_{eff} patterns as an effect of how training uses a fixed band.

The frozen-arm rows also document the mixed regime past $f_{\max} = 48$: train PSNR itself degrades ($42.9 \to 31.6$) while test PSNR partially recovers ($7.8 \to 10.4$). We report this regime as measured and leave it uninterpreted. Figure 5 plots the same sweep.

B Completed-grid tables

Table 3 collects the per-setting optima behind Figure 1. Tables 4 and 5 report mean test PSNR and mean B_{eff} for every candidate and depth on the completed grids, including the $f_{\max} = 20$ column. That column behaves consistently across depths: pre-expansion at $L = 2$, in the shoulder at $L = 3$, and fully expanded at $L \geq 4$ with collapsed test PSNR.

The B_{eff} columns of Table 3 vary within each setting, for example from 24.8 to 36.3 on CelebA. Figure 6 shows what this residual dispersion measures. Indexed by B_{eff}, the near-optimal configurations of all four depths fall inside a common band, roughly 24 to 36 on CelebA and 23 to 53 on

Table 2: Frozen versus free arm on the dense CelebA $L = 3$ sweep (8 images $\times$ 3 seeds per cell). Δ is the absolute aggregate test-PSNR gap.

	frozen arm			free arm			
$f_{\max}$	test	train	B_{eff}	test	train	B_{eff}	Δ (dB)
4	18.26	18.40	19	17.96	18.11	20	0.30
8	19.29	19.54	20	19.35	19.60	20	0.06
12	19.80	20.11	21	19.90	20.21	21	0.10
16	21.61	22.47	29	21.55	22.42	29	0.06
20	19.60	26.37	75	19.76	26.34	73	0.16
24	13.67	31.92	153	13.65	31.91	154	0.01
32	9.39	41.09	356	9.38	41.17	358	0.01
48	7.80	42.90	831	7.73	42.85	835	0.07
64	8.11	38.82	905	8.16	38.84	906	0.05
128	10.35	31.60	942	10.33	31.86	943	0.03

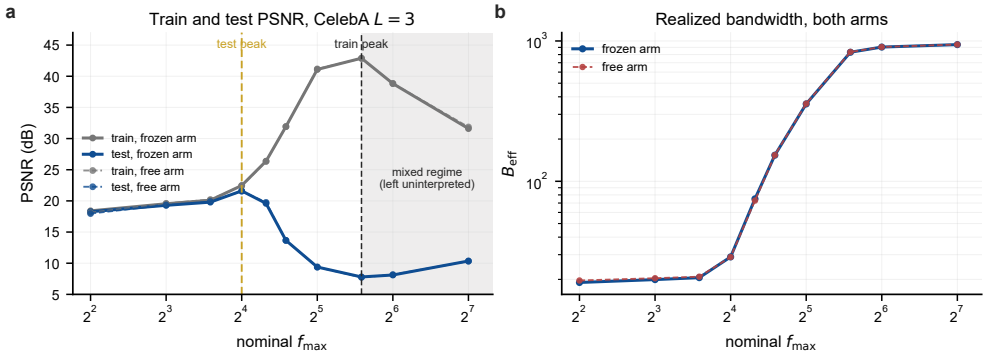


Figure 5: The control sweep of Table 2. (a) Train and test PSNR against $f_{\max}$ for both arms. Test PSNR peaks at $f_{\max} = 16$ and collapses past the boundary while train PSNR keeps rising to $f_{\max} = 48$. Past 48 (shaded) train PSNR itself degrades and test PSNR partially recovers. (b) Realized bandwidth in both arms. The frozen and free curves coincide. The expansion is therefore not attributable to drift of the activation frequencies.

Kodak, the near-optimal bands of Section 4.2. The band is the union of near-optimal B_{eff} values, not a region of uniformly good PSNR. On Kodak, $L = 4$ at $f_{\max} = 8$ realizes $B_{\text{eff}} = 30$ inside the band and still pays 1.97 dB against the depth optimum. Among the near-optimal candidates themselves the argmax is decided by tenths of a decibel, and a finer grid relocates it: the refined $L = 3$ sweep moves the CelebA optimum from $B_{\text{eff}} = 29.0$ to 41.3 for a gain of 0.40 dB. The dispersion of the optimum in Table 3 therefore reflects how many candidates tie within 0.5 dB on the evaluated grids.

C SIREN TABLES

Table 6 reports the SIREN sweeps. The $L = 3$ block uses 3 seeds, the depth block 2 seeds, 8 CelebA images throughout. The realized bandwidth grows steeply but gradually in ω_0 at every depth, and the largest adjacent-candidate ratio stays between 2.0 and 3.0 (Appendix D).

D SELECTOR ROBUSTNESS

Table 7 summarizes the checks of this appendix. Each row varies one ingredient of the selection pipeline and records what moved and whether any selection changed. Figure 7 plots the per-image regret distributions.

B_{sig} estimators. Three estimators give different absolute scales but identical selections. Full-image (oracle diagnostic only): $B_{\text{sig}} = 60.4/120.8$ (CelebA/Kodak), $\rho = 0.54$. Train-only linear (the

Table 3: Depth invariants on the completed grids. The nominal optimum moves with depth in every setting, while the realized bandwidth of the best configuration stays within a narrow band per setting. B_{eff} values are means over images and seeds at the optimum.

	CelebA (bank)		Kodak (bank)		CelebA (SIREN)	
L	$f^*_{\max}$	B_{eff} at $f^*_{\max}$	$f^*_{\max}$	B_{eff} at $f^*_{\max}$	ω^*_0	B_{eff} at ω^*_0
2	32	36.3	32	53.5	16	40.3
3	16	29.0	16	40.0	12	47.8
4	12	24.8	12	26.3	8	42.4
5	3	27.6	12	37.4	4	28.8

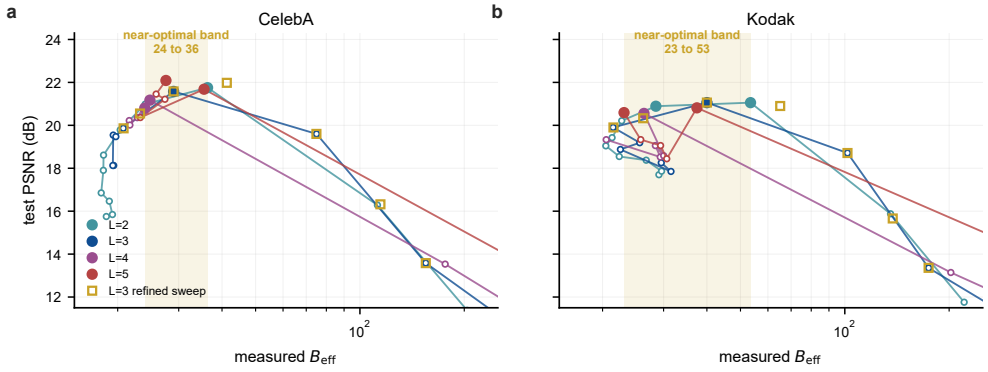


Figure 6: Test PSNR indexed by measured B_{eff} on the completed grids, a zoom of Figure 2b. Filled markers are the per-depth 0.5 dB near-optimal sets, open markers the remaining candidates, and the shaded band spans the union of the near-optimal B_{eff} values, 24–36 on CelebA and 23–53 on Kodak. Gold squares are the refined $L = 3$ sweep of Appendix H. The near-optimal points of all four depths fall inside the band, while the open markers inside it show that a candidate in the band is not necessarily near-optimal.

protocol): $B_{\text{sig}} = 42.0/81.5$, $\rho = 0.78$. Train-only nearest-neighbor: $B_{\text{sig}} = 73.4/140.2$, $\rho = 0.45$. Under all three, every selection in the four original protocols is identical. The cross-dataset B_{sig} ratio is $2.00/1.94/1.91$ under the three estimators, which is why calibrating with the same estimator compensates for its scale bias in these settings.

Per-image B_{sig} (train-linear / full-image). CelebA: 34.6–53.5 / 37.8–84.9 across the eight images. Kodak: 41.9–113.5 / 67.5–158.6. The dataset-level separation used by the selector is much larger than the within-dataset spread of the means across estimation masks reported below.

Estimation masks. Re-estimating B_{sig} with five alternative 25% coordinate masks changes the dataset-mean train-only B_{sig} by less than 2.5% (CelebA 41.4–42.4, Kodak 79.4–81.3) and changes no selection in the four original protocols, including the source-map selection of $f_{\max} = 20$ at Kodak $L = 3$.

Energy threshold. At 90% the Kodak own-map protocol attains 0.00 dB regret at all four depths, because its $L = 4$ selection moves from $f_{\max} = 8$ (regret 1.97 dB at 95%) to a near-optimal candidate. Every other selection is unchanged. The source-map shoulder failure persists (+2.35 dB at $L = 3$) and is therefore not a threshold artifact. At 99% the radius is dominated by the expanded tail and performance degrades broadly (own-map regrets 1.5–2.7 dB).

Boundary sensitivity. Under the 2× criterion the boundaries are $[32, 16, 12, 12]$ on both datasets. Under 3×, the $L = 3$ boundary disappears on both datasets, because the $f_{\max} = 20$ candidate splits the former single jump into two adjacent ratios of 2.58 and 2.06 (CelebA) and 2.54 and 1.71 (Kodak). Maximum adjacent ratios per depth: CelebA $3.09/2.58/7.11/20.47$, Kodak $2.54/2.54/7.68/20.88$, SIREN $2.98/2.09/2.05/2.15$. SIREN triggers the 2× rule only at scattered positions ($\omega_0 = 64/32/32/4$) and the 3× rule nowhere.

Table 4: CelebA completed grid (frozen arm). Cell = test PSNR (dB) / B_{eff}.

$f_{\max}$	$L = 2$	$L = 3$	$L = 4$	$L = 5$
3	15.8 / 19	18.1 / 19	20.2 / 22	22.1 / 28
4	15.8 / 19	18.1 / 19	20.0 / 22	21.5 / 26
6	16.5 / 19	19.5 / 19	20.8 / 24	21.2 / 27
8	16.9 / 18	19.5 / 20	20.4 / 23	20.4 / 23
12	17.9 / 18	19.9 / 21	21.2 / 25	21.7 / 35
16	18.6 / 18	21.6 / 29	13.5 / 176	10.1 / 727
20	19.8 / 20	19.6 / 75	9.6 / 442	10.2 / 927
24	21.0 / 24	13.6 / 155	9.3 / 758	10.6 / 952
32	21.8 / 36	9.3 / 360	9.9 / 911	11.0 / 962
48	16.3 / 112	7.8 / 831	10.7 / 955	11.1 / 964
64	11.4 / 204	8.1 / 904	11.0 / 962	11.0 / 963

Table 5: Kodak completed grid (frozen arm). Cell = test PSNR (dB) / B_{eff}.

$f_{\max}$	$L = 2$	$L = 3$	$L = 4$	$L = 5$
3	17.7 / 29	18.3 / 30	19.3 / 20	20.6 / 23
4	17.9 / 30	17.9 / 32	18.5 / 29	19.3 / 26
6	18.4 / 27	18.9 / 23	19.1 / 28	19.1 / 29
8	18.5 / 22	19.2 / 26	18.6 / 30	18.4 / 31
12	19.0 / 20	19.9 / 21	20.6 / 26	20.8 / 37
16	19.4 / 21	21.1 / 40	13.1 / 202	11.6 / 782
20	20.2 / 23	18.7 / 102	10.4 / 482	12.9 / 936
24	20.9 / 29	13.4 / 175	10.7 / 787	13.9 / 954
32	21.1 / 53	10.0 / 389	12.2 / 920	14.9 / 961
48	15.9 / 136	9.5 / 846	14.2 / 957	15.1 / 964
64	11.8 / 221	10.0 / 910	14.8 / 962	15.0 / 964

Coordinate splits. All main results use the fixed seed-123 coordinate permutation. We retrain CelebA under five alternative 25% splits (seeds $\{7, 17, 27, 37, 47\}$, network initialization fixed at seed 0) on the sub-grid $f_{\max} \in \{8, 12, 16, 20, 24, 32\}$ at $L = 3$ and $\{3, 4, 6, 8, 12, 16\}$ at $L = 5$ (480 fits), estimating B_{sig} from each split's own training coordinates. The structure remains unchanged. At $L = 3$ the optimum is $f^*_{\max} = 16$ and the selector regret is 0.00 dB under every split. At $L = 5$ the optimum is $f^*_{\max} = 3$, the sub-grid $2\times$ boundary sits at 12, and the selector picks 12 under every split, the plateau behavior of Section 4.2. Four of the five selections remain within the 0.5 dB near-optimal set, and the fifth exceeds it by 0.04 dB (regret 0.27–0.54 dB). The per-split B_{sig} stays within 41.4–42.4. These are one-dataset, sub-grid checks, and they support only split robustness.

Per-image regret. The selector is calibrated and evaluated at the dataset level. Against each image's own oracle on the evaluated grid, the dataset-level selection has median regret 0.0–0.4 dB across the seven held-out protocols, and interquartile ranges within $[0.0, 1.2]$ dB. The tails are real: the 95th percentiles span 0.77–3.03 dB, and the worst single image pays 4.43 dB, on one image of the held-out CelebA cohort. The joint Kodak-SIREN protocol is the tightest, with median 0.00 and 95th percentile 0.77 dB.

Bootstrap. The image-level bootstrap (2,000 resamples) recomputes test-PSNR means, the B_{eff} maps, B_{sig}, ρ, the grid optimum, and the selection per resample, with CelebA and SIREN paired on the same image resample. This full recomputation applies to the four original protocols. For the three frozen-protocol confirmations, ρ stays fixed in every resample, because those protocols never re-fit it (Appendix H). Aggregate $f^*_{\max}$ reproduction rates: CelebA $1.00/1.00/0.82/0.70$ and Kodak $0.92/1.00/0.999/0.75$ for $L = 2/3/4/5$.

Cost of selection. The three ways to set the knob on a new dataset differ in what must be trained there. A direct sweep trains every candidate at every depth, $11 \times 4 \times 8 \times 2 = 704$ fits on the completed Kodak grid, and reaches the grid optimum by construction. The own-map protocol trains the same 704 fits to build the map and then selects without looking at test PSNR, at a mean regret of 0.49 dB. The source-map protocol trains only the selected candidate, 64 fits, at a mean regret of

Table 6: SIREN. Left: $L = 3$ full sweep. Right: depth sweep, cell = test PSNR / B_{eff}.

	$L = 3$			depth sweep			
ω_0	test	train	B_{eff}	$L = 2$	$L = 3$	$L = 4$	$L = 5$
4	19.8	20.0	21	17.2 / 18	19.7 / 20	21.2 / 25	22.1 / 29
8	22.5	24.2	33	20.3 / 24	22.5 / 32	22.6 / 42	21.6 / 62
12	22.8	28.0	49	22.2 / 31	22.8 / 48	20.7 / 75	16.9 / 120
16	20.1	31.5	76	22.6 / 40	20.2 / 75	14.0 / 131	10.0 / 195
20	15.4	36.5	115	21.6 / 53	15.6 / 112	9.4 / 179	7.3 / 286
24	11.2	47.1	146	19.8 / 71	11.4 / 144	7.3 / 247	6.3 / 443
32	7.5	63.0	209	14.1 / 112	7.6 / 206	6.0 / 464	6.1 / 874
48	5.9	55.4	436	7.9 / 175	5.9 / 430	6.2 / 953	9.2 / 964
64	5.7	50.3	765	6.4 / 251	5.7 / 759	6.8 / 963	7.8 / 964
128	6.6	38.7	964	5.7 / 748	6.6 / 964	7.9 / 963	7.9 / 964

Table 7: Selector robustness at a glance. Details and full numbers in the paragraphs below.

Check	Variation	What moved	Selections and regret
B_{sig} estimator	full-image, train-linear (protocol), train-NN	B_{sig} $60.4/42.0/73.4$ on CelebA, ρ $0.54/0.78/0.45$	every selection identical
Estimation mask	five alternative 25% masks	dataset-mean B_{sig} within 2.5%	no selection changes in the four original protocols
Energy threshold	90%, 95% (protocol), 99%	90%: Kodak own-map selection at $L = 4$ changes. 99%: radius dominated by expanded tail	90%: own-map 0.00 dB at all four Kodak depths, shoulder miss persists. 99%: own-map 1.5–2.7 dB
Boundary criterion	$2\times$ (protocol), $3\times$	$3\times$ removes the $L = 3$ boundary on both datasets	$2\times$: $[32, 16, 12, 12]$ on both datasets
Coordinate split	five 25% splits, 480 fits	B_{sig} 41.4–42.4	$L = 3$: 0.00 dB on every split. $L = 5$: 0.27–0.54 dB
Per-image regret	seven held-out protocols	median 0.0–0.4 dB, 95th percentile 0.77–3.03 dB	worst single image 4.43 dB
Bootstrap	2,000 image resamples, full pipeline	argmax reproduced $1.00/1.00/0.82/0.70$ (CelebA), $0.92/1.00/0.999/0.75$ (Kodak)	CIs of Table 1

0.59 dB and 2.35 dB at the shoulder. The saving is a factor of 11 in target-side fits, and the price is paid entirely near the expansion shoulder.

E RAW VERSUS CLIPPED SPECTRA

PSNR uses predictions clipped to $[0, 1]$, while B_{eff} uses raw outputs. Across all 142 fits of the grid-completion run, which record both spectra, the median relative difference in B_{eff} is 0.02%, the 95th percentile 6.5%, and the maximum 7.1%. The difference is negligible in the pre-boundary regime (e.g. $f_{\max} = 4$: 19.9 vs 19.9; $f_{\max} = 16$: 31.4 vs 31.4) and small in the fully expanded regime (e.g. $f_{\max} = 48$: 825 vs 838), where outputs escape $[0, 1]$ and clipping adds mild high-frequency content. These checks indicate that the raw-output convention is not driving the reported grid-completion behavior.

F FOURIER-FEATURE CONTROL

A random-Fourier-feature network (Tancik et al., 2020) with a ReLU MLP is trained on the same task family across the same depths, sweeping the input scale σ. The ReLU width is solved per depth to match the bank's trainable parameter count within 0.4% (full numbers in Appendix H). The

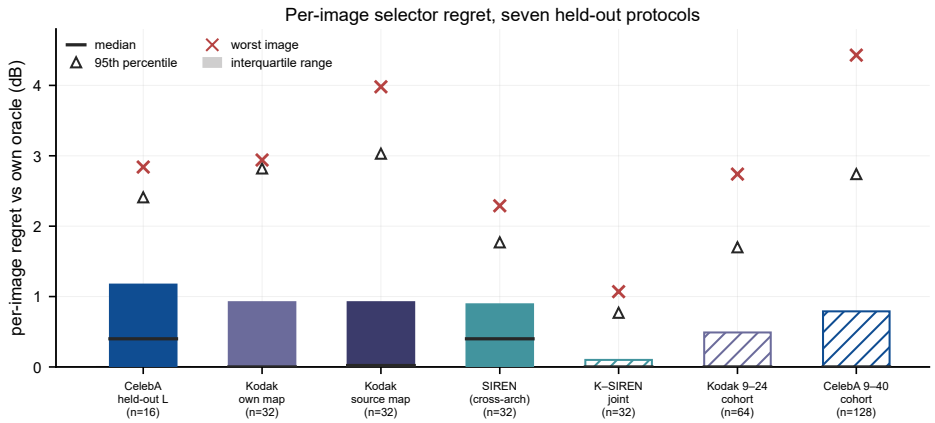


Figure 7: Per-image selector regret against each image's own grid oracle for the seven held-out protocols, drawn from the summary statistics above: interquartile range (box), median (bar), 95th percentile (triangle), and worst image (cross). Hatched boxes are the frozen-ρ confirmations. Medians sit at or near zero in every protocol, while worst-case regrets reach 4.43 dB. Here n counts image–depth pairs.

aggregate optimum is $\sigma^* = 1/1/1/0.5$, and $\sigma = 1$ remains within the 0.5 dB near-optimal set at every depth. No bank-like depth-dependent shift appears in this control.

G Implementation details and reproduction

Table 8 lists the default training and measurement settings of the main sinusoidal sweeps. Exceptions, such as the parameter-matched ReLU widths and the alternative coordinate splits, are specified in Appendices D and H. Table 9 lists the run blocks and their fit counts. All analyses are post-processing of archived run records. The locked summary statistics are reproduced by the verification script, and the appendix tables are direct aggregates of the corresponding raw JSON records.

Table 8: Default training and measurement settings of the main sinusoidal sweeps. Exceptions are specified in Appendices D and H.

Item	Setting
Width	256 at every depth
Bank activation	$\tau = 8$ frequencies per unit, $\mathrm{geomspace}(2, f_{\max}, 8)$, frozen in the main arm
SIREN initialization	first layer $U(-1/n_{\text{in}}, 1/n_{\text{in}})$, hidden layers $U(\pm\sqrt{6/n_{\text{in}}}/\omega_0)$ (Sitzmann et al., 2020)
Optimizer	Adam, learning rate 10^{-3}, exponential decay by factor 0.05 over 800 iterations
Batch and loss	full batch on the 25% training coordinates, mean-squared error
Model selection	model after iteration 800, no early stopping, held-out pixels unseen during training
Images	128×128 RGB in $[0, 1]$, coordinates in $[-1, 1]^2$, split seed 123
Spectrum	512^2 evaluation grid, channel-summed 2D power, DC removed, radial cumulative energy
B_{eff} threshold	95% of spectral energy (sensitivity in Appendix D)

H Confirmation experiments

Six confirmation and refinement experiment blocks were executed after the analysis definitions of Section 3.4 were frozen. No threshold, boundary definition, or calibration ratio was reselected for

Table 9: Run blocks. Grid dimensions are knobs × depths × images × seeds unless stated.

Block	Grid	Fits
CelebA depth grid	$10 \times 4 \times 8 \times 3$	960
Kodak depth grid	$10 \times 4 \times 8 \times 2$	640
Dense CelebA $L = 3$ sweep, two arms	$10 \times 2 \times 8 \times 3$	480
SIREN $L = 3$ sweep	$10 \times 8 \times 3$	240
SIREN depth grid	$10 \times 4 \times 8 \times 2$	640
Grid completion		142
Confirmation runs (Appendix H)	$640 + 512 + 96 + 704$	1,952
Refinement and CelebA cohort (Appendix H)	$96 + 1{,}408$	1,504
Split robustness (Appendix D)	five splits, sub-grids	480
Total instrumented fits		7,038
Layerwise measurements	12 configurations	

them: $\rho = 0.78$, the seed-123 split, and the 0.5 dB near-optimal tolerance carry over unchanged. The shoulder refinement adds one statistic, the grid-normalized slope, which was fixed before the run because adjacent-ratio rules are not comparable across grid densities. For the three selector protocols we additionally report image-level bootstrap intervals (2,000 resamples) that recompute the maps, B_{sig}, and the grid optimum per resample while keeping ρ fixed. Table 10 summarizes the six blocks.

Table 10: Confirmation and refinement blocks, all run after the analysis definitions were frozen. Each row names what the block holds out or refines and the headline result. Full numbers in the paragraphs below.

Block	Setup (fits)	Tests	Result
Kodak-SIREN grid	8 Kodak images, 2 seeds (640)	dataset and architecture together	$\omega_0^* = 12/8/8/4$, spans $4.0\times$ vs $1.77\times$, $3\times$ rule fires nowhere, frozen-ρ regret 0.06 dB
Parameter-matched Fourier-feature control	ReLU widths $160/201/217/226$, 8 CelebA images, 2 seeds (512)	depth dependence in a parameter-matched input-encoding control	$\sigma^* = 1/1/1/0.5$, $\sigma = 1$ near-optimal at every depth
Multi-image layer-wise	4 images, 2 seeds, both datasets, $L \in \{3, 5\}$, three regimes (96)	replication of the late-layer amplification	final layer $\geq 2\times$ first layer in $32/32$ past-boundary and $0/32$ pre-boundary pairs
Held-out Kodak cohort	images 9–24, one seed (704)	fresh images	boundary $[32, 16, 12, 12]$, spans $10.67\times$ vs $1.94\times$, regret 0.08 dB
Held-out CelebA cohort	images 9–40, one seed (1,408)	fresh images of the source dataset	boundary $[32, 16, 12, 12]$, spans $10.67\times$ vs $1.57\times$, regret 0.12 dB
Shoulder refinement	$f_{\max} \in \{14, 18, 22\}$ at $L = 3$, both datasets (96)	grid dependence of the boundary	peak slope 5.66 (CelebA) and 4.25 (Kodak) on $[18, 20]$, refined near-optimal set $\{16, 18\}$

Kodak-SIREN grid. The SIREN sweep of Section 3.2 on the eight Kodak images (2 seeds, 640 fits). The nominal optimum moves, $\omega_0^* = 12/8/8/4$ for $L = 2/3/4/5$, while near-optimal values span $4.0\times$ and the corresponding realized B_{eff} spans $1.77\times$ (band 39.5–70.0). The $3\times$ rule fires nowhere and the largest adjacent-step ratios per depth are $2.95/2.07/2.19/2.43$, matching the graded expansion of Section 5.2. Selecting with the frozen ρ on this map gives per-depth regrets $0.09/0.16/0.00/0.00$ dB: mean point estimate 0.06 dB, 95% bootstrap CI $[0.00, 0.37]$, per-depth maximum 0.16 dB.

Parameter-matched Fourier-feature control. ReLU widths $160/201/217/226$ for $L = 2/3/4/5$ match the frozen-arm bank's trainable parameter count within 0.4% (mapping size 128, with a non-

trainable Gaussian B sampled from each run's seed). On the σ grid $\{0.5, 1, 2, 4, 8, 16, 32, 64\}$ with 8 CelebA images and 2 seeds (512 fits), the aggregate optimum is $\sigma^* = 1/1/1/0.5$ and every near-optimal set contains $\sigma = 1$. The realized B_{eff} at the optimum stays within 31.5–39.2. In this parameter-matched control, no bank-like depth-dependent shift appears: $\sigma = 1$ remains within the 0.5 dB near-optimal tolerance at every depth.

Multi-image layerwise. Four images, two seeds, both datasets, $L \in \{3, 5\}$, and three regimes per depth: pre-boundary $f_{\max} = 8$, shoulder 20 ($L = 3$) or 16 ($L = 5$), post-boundary 32 (96 fits, frozen arm). Median per-layer B_{eff} sequences at $L = 5$ on CelebA: pre $38 \to 38 \to 28 \to 20 \to 22$, shoulder $38 \to 25 \to 53 \to 249 \to 622$, post $25 \to 51 \to 424 \to 709 \to 721$. At $L = 5$, the Kodak medians match closely in the shoulder and post-boundary regimes and stay flat at 35–38 in the pre-boundary regime. A final layer with at least twice the first layer's B_{eff} appears in $32/32$ past-boundary image-seed pairs, $0/32$ pre-boundary pairs, and $16/16$ shoulder pairs at $L = 5$ versus $0/16$ at $L = 3$, where the shoulder elevation (66 versus 38 in the final layer) stays below the $2\times$ counting threshold.

Held-out Kodak cohort. Kodak images 9–24, completed grid, frozen arm, one seed (704 fits). The $2\times$ boundary is $[32, 16, 12, 12]$, as on the original subset. Near-optimal nominal values span $10.67\times$ while realized B_{eff} spans $1.94\times$ (band 24.0–46.6). Maximum adjacent ratios are $2.78/2.81/7.97/21.81$. The deep-end argmax lands at $f^*_{\max} = 6/3$ for $L = 4/5$, with near-optimal sets containing 12, which is the plateau degeneracy of Section 4.2. Selecting with the frozen ρ gives per-depth regrets $0.00/0.00/0.01/0.31$ dB: mean point estimate 0.08 dB, 95% bootstrap CI $[0.00, 0.39]$, per-depth maximum 0.31 dB.

Held-out CelebA cohort. CelebA images 9–40, completed grid, frozen arm, one seed (1,408 fits). The $2\times$ boundary is $[32, 16, 12, 12]$, as on the original subset. Near-optimal nominal values span $10.67\times$ while realized B_{eff} spans $1.57\times$ (band 24.0–37.7). Maximum adjacent ratios are $3.10/2.63/7.77/24.56$. The deep-end near-optimal sets contain several nominal values, $\{6, 12\}$ at $L = 4$ and $\{3, 12\}$ at $L = 5$, the plateau degeneracy of Section 4.2. The cohort's train-only B_{sig} mean is 41.2, compared with 42.0 on the original eight images. Selecting with the frozen ρ gives per-depth regrets $0.00/0.00/0.00/0.48$ dB: mean point estimate 0.12 dB, 95% bootstrap CI $[0.00, 0.45]$, per-depth maximum 0.48 dB.

Shoulder refinement. Adding $f_{\max} \in \{14, 18, 22\}$ at $L = 3$ on both datasets (8 images, 2 seeds, 96 fits) refines the $L = 3$ curve to seven points between 12 and 24. The grid-normalized slopes $s_i = \Delta \log B_{\text{eff}} / \Delta \log f_{\max}$ per adjacent pair are $0.72/1.68/3.00/5.66/4.46/3.46$ on CelebA and $1.30/3.17/4.13/4.25/3.15/2.74$ on Kodak, peaking on the $[18, 20]$ pair in both datasets. On the refined curve no adjacent B_{eff} ratio reaches $2\times$, because the inserted candidates split the former coarse-grid jump, which is the reason this block reports slopes. The aggregate test optimum moves to $f_{\max} = 18$ on CelebA (21.98 versus 21.58 dB at 16) and stays at 16 on Kodak (21.06 versus 20.90 dB at 18). The refined $L = 3$ near-optimal set is $\{16, 18\}$ on both datasets, ending at the near endpoint of the peak-slope interval $[18, 20]$. Figure 8 plots this block and the multi-image layerwise block.

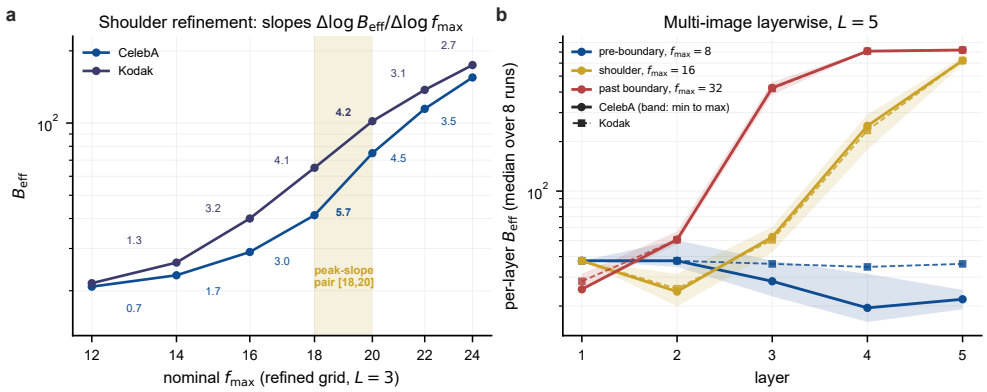


Figure 8: Two confirmation blocks. (a) The refined $L = 3$ curve on both datasets with the grid-normalized slope of each adjacent pair. The peak slope falls on the $[18, 20]$ pair (shaded) for CelebA and Kodak alike. (b) Per-layer B_{eff} at $L = 5$ in the three regimes of the multi-image layerwise block, median over four images and two seeds (CelebA solid with min–max band, Kodak dashed). The pre-boundary stack stays flat, the shoulder and past-boundary stacks amplify in the late layers.