NNGPResNet: IID and Correlated-Layer Bounds

Anthony Andiles

February 12, 2026

Abstract

This note gives complete mathematical proofs for the IID and correlated NNGPResNet results formalized in Lean. All constants are chosen to match the Lean theorems. A final section maps every definition and theorem in this note to the exact Lean declaration and file path.

1 Setup and Notation

Fix integers $m, r, L \in \mathbb{N}$ with $L \geq 1$, $m \geq 1$, and define $rm := rm$. For $x \in \mathbb{R}$, define $\operatorname{sign}(x) = 1$ if $x \geq 0$, and -1 if $x < 0$. For vectors, $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $\| \cdot \|$ is the Euclidean norm. The unit sphere is

$$S^{m-1} := \{u \in \mathbb{R}^m : \|u\| = 1\}.$$

For a centered Gaussian random variable $G \sim \mathcal{N}(0, \sigma^2)$ we use:

$$\mathbb{P}(|G| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\sigma^2}\right), \qquad \mathbb{E}[G^2] = \sigma^2.$$

2 IID NNGP Model

2.1 Model definition

For each $\ell \in \{1, \dots, L\}$, $i \in \{1, \dots, m\}$, $j \in \{1, \dots, rm\}$:

- $B_{\ell ij} \sim \mathcal{N}(0, 1)$ i.i.d.,
- $z_{\ell j} \sim \mathcal{N}(0, 1)$ i.i.d., independent of all B,
- $a_{\ell j} := \operatorname{sign}(z_{\ell j}) \in \{\pm 1\}$.

Define layer increment and probes:

$$\Delta h_\ell(i) := \frac{1}{\sqrt{L}} \frac{1}{\sqrt{rm}} \sum_{j=1}^{rm} B_{\ell ij} a_{\ell j}, \quad g_\ell(u) := \langle u, \Delta h_\ell \rangle,$$

$$M_u := \frac{1}{L} \sum_{\ell=1}^{L} g_\ell(u), \qquad W := \frac{1}{L} \sum_{\ell=1}^{L} \Delta h_\ell, \qquad X(u) := \langle u, W \rangle.$$

Clearly $X(u) = M_u$ for every u.

2.2 IID-T1: exact scalar law, tail, second moment

Theorem 2.1 (IID-T1 Gaussian law)**.** *For each fixed $u \in S^{m-1}$,*

$$M_u \sim \mathcal{N}\left(0, \frac{1}{L^2}\right).$$

Proof. Fix $u \in S^{m-1}$. For one layer ℓ,

$$g_\ell(u) = \frac{1}{\sqrt{L}\sqrt{rm}} \sum_{i=1}^{m} \sum_{j=1}^{rm} u_i a_{\ell j} B_{\ell ij}.$$

Condition on $\{a_{\ell j}\}_{j=1}^{rm}$. Then $g_\ell(u)$ is a linear combination of independent centered Gaussians, hence Gaussian with mean 0 and variance

$$\mathrm{Var}(g_\ell(u) \mid a) = \frac{1}{Lrm} \sum_{i=1}^{m} \sum_{j=1}^{rm} u_i^2 a_{\ell j}^2 = \frac{1}{Lrm} \left(\sum_{i=1}^{m} u_i^2 \right) \left(\sum_{j=1}^{rm} 1 \right) = \frac{\|u\|^2}{L} = \frac{1}{L}.$$

So unconditionally $g_\ell(u) \sim \mathcal{N}(0, 1/L)$. For different ℓ, variables $g_\ell(u)$ depend on disjoint Gaussian coordinates $(B_{\ell ij}, z_{\ell j})$, so they are independent. Hence

$$M_u = \frac{1}{L} \sum_{\ell=1}^{L} g_\ell(u)$$

is centered Gaussian with variance

$$\mathrm{Var}(M_u) = \frac{1}{L^2} \sum_{\ell=1}^{L} \mathrm{Var}(g_\ell(u)) = \frac{1}{L^2} \cdot L \cdot \frac{1}{L} = \frac{1}{L^2}.$$

Therefore $M_u \sim \mathcal{N}(0, 1/L^2)$. □

Corollary 2.2 (IID-T1 tail)**.** *For each $u \in S^{m-1}$ and each $t > 0$,*

$$\mathbb{P}(|M_u| \geq t) \leq 2 \exp\left(-\frac{L^2 t^2}{2} \right).$$

Proof. Apply the standard Gaussian tail inequality to Theorem **??** with variance $\sigma^2 = 1/L^2$. □

Corollary 2.3 (IID-T1 second moment)**.** *For each $u \in S^{m-1}$,*

$$\mathbb{E}[M_u^2] = \frac{1}{L^2}.$$

Proof. A centered Gaussian has second moment equal to variance, and Theorem **??** gives variance $1/L^2$. □

2.3 IID-T2: expected supremum via Dudley

Fix $u_0 \in S^{m-1}$ and define centered process

$$\widetilde{X}(u) := X(u) - X(u_0), \qquad u \in S^{m-1}.$$

Lemma 2.4 (Increment metric)**.** *For all $u, v \in S^{m-1}$,*

$$X(u) - X(v) \sim \mathcal{N}\left(0, \frac{\|u-v\|^2}{L^2}\right),$$

so the canonical metric is

$$d(u,v) := \sqrt{\mathbb{E}[(X(u) - X(v))^2]} = \frac{1}{L}\|u - v\|.$$

Proof. $X(u) - X(v) = M_{u-v}$. Repeating the computation in Theorem **??** for a general vector $w \in \mathbb{R}^m$ gives $M_w \sim \mathcal{N}(0, \|w\|^2/L^2)$. Set $w = u - v$. □

Lemma 2.5 (Entropy integral bound with explicit constant)**.** *If $m \geq 1$, then*

$$\mathrm{EntInt}(S^{m-1}, 2) := \int_0^2 \sqrt{\log N(S^{m-1}, \varepsilon)}\, d\varepsilon \leq 4\sqrt{m}.$$

Proof. Let $B_2^m := \{x \in \mathbb{R}^m : \|x\| \leq 1\}$. Since $S^{m-1} \subset B_2^m$, monotonicity of covering numbers gives

$$N(S^{m-1}, \varepsilon) \leq N(B_2^m, \varepsilon).$$

For $0 < \varepsilon \leq 2$, the standard volumetric estimate gives

$$N(B_2^m, \varepsilon) \leq \left(1 + \frac{2}{\varepsilon}\right)^m \leq \left(\frac{3}{\varepsilon}\right)^m.$$

Hence

$$\log N(S^{m-1}, \varepsilon) \leq m \log\left(\frac{3}{\varepsilon}\right)$$

and therefore

$$\mathrm{EntInt}(S^{m-1}, 2) \leq \sqrt{m} \int_0^2 \sqrt{\log\left(\frac{3}{\varepsilon}\right)}\, d\varepsilon.$$

Set

$$I := \int_0^2 \sqrt{\log\left(\frac{3}{\varepsilon}\right)}\, d\varepsilon.$$

Use substitution $\varepsilon = 3e^{-s}$, $d\varepsilon = -3e^{-s} ds$. As $\varepsilon \downarrow 0$, $s \to \infty$; as $\varepsilon = 2$, $s = \log(3/2)$. Thus

$$I = 3 \int_{\log(3/2)}^{\infty} e^{-s}\sqrt{s}\, ds \leq 3 \int_0^{\infty} e^{-s}\sqrt{s}\, ds = 3\,\Gamma\left(\frac{3}{2}\right) = \frac{3\sqrt{\pi}}{2} < 4.$$

So $\mathrm{EntInt}(S^{m-1}, 2) \leq 4\sqrt{m}$. □

Theorem 2.6 (IID-T2 centered Dudley bound)**.** *Assume $m \geq 1$ and $L \geq 1$. Then*

$$\mathbb{E}\Big[\sup_{u \in S^{m-1}} \widetilde{X}(u)\Big] \leq (12\sqrt{2})\frac{1}{L}\,\mathrm{EntInt}(S^{m-1}, 2) \leq (48\sqrt{2})\frac{\sqrt{m}}{L}.$$

Proof. By Lemma **??**, increments are sub-Gaussian with scale $1/L$ in Euclidean distance. Apply Dudley's theorem with the explicit constant $12\sqrt{2}$ used in the Lean development:

$$\mathbb{E} \sup_{u \in S^{m-1}} \widetilde{X}(u) \leq (12\sqrt{2}) \frac{1}{L} \operatorname{EntInt}(S^{m-1}, 2).$$

Then use Lemma **??**. □

Theorem 2.7 (IID-T2 uncentered bound)**.** *Assume $m \geq 1$ and $L \geq 1$. Then*

$$\mathbb{E}\Big[\sup_{u \in S^{m-1}} X(u) \Big] \leq (48\sqrt{2}) \frac{\sqrt{m}}{L}.$$

Proof. For every sample,

$$\sup_{u \in S^{m-1}} X(u) = \sup_{u \in S^{m-1}} (X(u) - X(u_0)) + X(u_0)$$

because adding the constant $-X(u_0)$ shifts all values equally. Integrate both sides:

$$\mathbb{E} \sup_u X(u) = \mathbb{E} \sup_u \widetilde{X}(u) + \mathbb{E}[X(u_0)].$$

By Theorem **??** with unit vector u_0, $X(u_0) = M_{u_0} \sim \mathcal{N}(0, 1/L^2)$, so $\mathbb{E}[X(u_0)] = 0$. Hence

$$\mathbb{E} \sup_u X(u) = \mathbb{E} \sup_u \widetilde{X}(u) \leq (48\sqrt{2}) \frac{\sqrt{m}}{L}$$

by Theorem **??**. □

2.4 IID-T3: high-probability supremum bound (Lean-exported form)

Theorem 2.8 (IID-T3)**.** *Assume $m \geq 1$, $L \geq 1$, and $t > 0$. Then*

$$\mathbb{P}\left(\mathbb{E} \sup_{u \in S^{m-1}} X(u) + t \leq \sup_{u \in S^{m-1}} X(u) \right) \leq 2m \exp\left(-\frac{L^2}{2} \left(\frac{t}{\sqrt{m}} \right)^2 \right).$$

Proof. Let $S(\omega) := \sup_{u \in S^{m-1}} X(u, \omega)$. Since $-u \in S^{m-1}$ whenever $u \in S^{m-1}$,

$$S(\omega) = \sup_{u \in S^{m-1}} \langle u, W(\omega) \rangle \geq 0,$$

so $\mathbb{E}[S] \geq 0$. Therefore

$$\{\mathbb{E}[S] + t \leq S\} \subseteq \{t \leq S\}.$$

Now prove

$$\{t \leq S\} \subseteq \bigcup_{i=1}^{m} \left\{ |W_i| \geq \frac{t}{\sqrt{m}} \right\}.$$

Indeed, if all coordinates satisfy $|W_i| < t/\sqrt{m}$, then

$$\|W\|^2 = \sum_{i=1}^{m} W_i^2 < m \cdot \frac{t^2}{m} = t^2,$$

so $\|W\| < t$. Then for any unit u,

$$X(u) = \langle u, W \rangle \le \|u\| \, \|W\| = \|W\| < t,$$

thus $S < t$, contradiction. Hence by union bound,

$$\mathbb{P}(t \le S) \le \sum_{i=1}^{m} \mathbb{P}\left(|W_i| \ge \frac{t}{\sqrt{m}}\right).$$

For each coordinate basis vector $e_i \in S^{m-1}$, $W_i = M_{e_i}$, so by Corollary **??**,

$$\mathbb{P}\left(|W_i| \ge \frac{t}{\sqrt{m}}\right) \le 2 \exp\left(-\frac{L^2}{2}\left(\frac{t}{\sqrt{m}}\right)^2\right).$$

Summing over $i = 1, \dots, m$ gives the claimed bound. □

3 Correlated-Layer Model and Effective Depth

3.1 Model definition

Fix $A \in \mathbb{R}^{L \times L}$ and define

$$\rho(\ell, k) := \sum_{t=1}^{L} A_{\ell t} A_{kt}, \qquad \rho_{\text{sum}} := \sum_{\ell=1}^{L} \sum_{k=1}^{L} \rho(\ell, k),$$

$$\text{invDEff} := \frac{\rho_{\text{sum}}}{L^3}, \qquad D_{\text{eff}} := \frac{L^3}{\rho_{\text{sum}}} \quad (\rho_{\text{sum}} > 0).$$

Randomness in the Lean model:

- Latent field $Z_{tij} \sim \mathcal{N}(0, 1)$ i.i.d. over (t, i, j),
- Shared activation Gaussian $z_j \sim \mathcal{N}(0, 1)$ i.i.d. over j,
- $a_j := \text{sign}(z_j)$ is shared across layers.

Define correlated weights and probes:

$$B_{\ell ij} := \sum_{t=1}^{L} A_{\ell t} Z_{tij},$$

$$\Delta h_\ell(i) := \frac{1}{\sqrt{L}\sqrt{rm}} \sum_{j=1}^{rm} B_{\ell ij} a_j, \qquad g_\ell(u) := \langle u, \Delta h_\ell \rangle,$$

$$M_u^{\text{layerAvg}} := \frac{1}{L} \sum_{\ell=1}^{L} g_\ell(u).$$

The Lean development also uses a flattened linear-form version M_u^{flat} and proves

$$M_u^{\text{layerAvg}} = M_u^{\text{flat}}.$$

3.2 CORR-COV and CORR-VAR

Theorem 3.1 (CORR-COV)**.** *For every $u \in \mathbb{R}^m$ and $\ell, k \in \{1, \ldots, L\}$,*

$$\mathrm{Cov}(g_\ell(u), g_k(u)) = \frac{\rho(\ell, k)}{L} \|u\|^2.$$

Proof. The Lean development proves the exact variance identities

$$\mathrm{Var}(g_\ell(u)) = \frac{\rho(\ell, \ell)}{L}\|u\|^2,$$

$$\mathrm{Var}(g_\ell(u) + g_k(u)) = \frac{\rho(\ell, \ell) + \rho(k, k) + 2\rho(\ell, k)}{L}\|u\|^2.$$

Now use

$$\mathrm{Cov}(X, Y) = \frac{\mathrm{Var}(X + Y) - \mathrm{Var}(X) - \mathrm{Var}(Y)}{2}$$

with $X = g_\ell(u)$, $Y = g_k(u)$ to obtain

$$\mathrm{Cov}(g_\ell(u), g_k(u)) = \frac{1}{2L} \left(\rho(\ell, \ell) + \rho(k, k) + 2\rho(\ell, k) - \rho(\ell, \ell) - \rho(k, k)\right) \|u\|^2 = \frac{\rho(\ell, k)}{L}\|u\|^2.$$

□

Theorem 3.2 (CORR-VAR: covariance-sum identity)**.** *For every $u \in \mathbb{R}^m$,*

$$\mathrm{Var}(M_u^{\mathrm{layerAvg}}) = \frac{1}{L^2} \sum_{\ell=1}^{L} \sum_{k=1}^{L} \mathrm{Cov}(g_\ell(u), g_k(u)).$$

Hence,

$$\mathrm{Var}(M_u^{\mathrm{layerAvg}}) = \frac{\|u\|^2 \rho_{\mathrm{sum}}}{L^3} = \frac{\|u\|^2}{D_{\mathrm{eff}}} \quad (\rho_{\mathrm{sum}} > 0).$$

Proof. Since $M_u^{\mathrm{layerAvg}} = (1/L) \sum_\ell g_\ell(u)$,

$$\mathrm{Var}(M_u^{\mathrm{layerAvg}}) = \mathrm{Var}\left(\frac{1}{L} \sum_\ell g_\ell(u)\right) = \frac{1}{L^2} \sum_{\ell,k} \mathrm{Cov}(g_\ell(u), g_k(u)).$$

Substitute Theorem **??**:

$$\mathrm{Var}(M_u^{\mathrm{layerAvg}}) = \frac{1}{L^2} \sum_{\ell,k} \frac{\rho(\ell, k)}{L}\|u\|^2 = \frac{\|u\|^2}{L^3} \sum_{\ell,k} \rho(\ell, k) = \frac{\|u\|^2 \rho_{\mathrm{sum}}}{L^3}.$$

If $\rho_{\mathrm{sum}} > 0$, then $D_{\mathrm{eff}} = L^3/\rho_{\mathrm{sum}}$, so

$$\frac{\|u\|^2 \rho_{\mathrm{sum}}}{L^3} = \frac{\|u\|^2}{D_{\mathrm{eff}}}.$$

□

Theorem 3.3 (Correlated scalar Gaussian law, tail, MSE for unit vectors)**.** *Assume $u \in S^{m-1}$. Then*

$$M_u^{\mathrm{layerAvg}} \sim \mathcal{N}\Big(0, \frac{\rho_{\mathrm{sum}}}{L^3}\Big) = \mathcal{N}\Big(0, \frac{1}{D_{\mathrm{eff}}}\Big) \qquad (\rho_{\mathrm{sum}} > 0),$$

$$\mathbb{P}\left(|M_u^{\mathrm{layerAvg}}| \geq t\right) \leq 2\exp\left(-\frac{t^2 D_{\mathrm{eff}}}{2}\right), \qquad t > 0,$$

$$\mathbb{E}\left[(M_u^{\mathrm{layerAvg}})^2\right] = \frac{\rho_{\mathrm{sum}}}{L^3} = \frac{1}{D_{\mathrm{eff}}}.$$

Proof. The process M_u^{flat} is a finite linear form in independent Gaussian coordinates; therefore it is centered Gaussian with variance $\|u\|^2 \rho_{\mathrm{sum}}/L^3$. Using $\|u\| = 1$ and the bridge $M_u^{\mathrm{layerAvg}} = M_u^{\mathrm{flat}}$ gives the law. The tail and second moment are the standard centered Gaussian formulas with variance ρ_{sum}/L^3; when $\rho_{\mathrm{sum}} > 0$, rewrite as $1/D_{\mathrm{eff}}$. □

3.3 CORR-T2: expected supremum via Dudley

Define $X_{\mathrm{corr}}(u) := M_u^{\mathrm{layerAvg}}$ and centered process

$$\widetilde{X}_{\mathrm{corr}}(u) := X_{\mathrm{corr}}(u) - X_{\mathrm{corr}}(u_0), \qquad u \in S^{m-1}.$$

Theorem 3.4 (CORR-T2)**.** *Assume $m \geq 1$, $L \geq 1$, and $\rho_{\mathrm{sum}} > 0$. Then*

$$\mathbb{E} \sup_{u \in S^{m-1}} \widetilde{X}_{\mathrm{corr}}(u) \leq (48\sqrt{2}) \frac{\sqrt{m}}{\sqrt{D_{\mathrm{eff}}}}.$$

Proof. From the correlated increment law in Lean, increments are sub-Gaussian with metric

$$d_{\mathrm{corr}}(u, v) = \sqrt{\mathrm{invDEff}}\, \|u - v\| = \frac{\|u - v\|}{\sqrt{D_{\mathrm{eff}}}}.$$

Applying Dudley with constant $12\sqrt{2}$ gives

$$\mathbb{E} \sup_u \widetilde{X}_{\mathrm{corr}}(u) \leq (12\sqrt{2})\sqrt{\mathrm{invDEff}}\ \mathrm{EntInt}(S^{m-1}, 2).$$

By Lemma **??**,

$$\mathbb{E} \sup_u \widetilde{X}_{\mathrm{corr}}(u) \leq (12\sqrt{2})\sqrt{\mathrm{invDEff}} \cdot 4\sqrt{m} = (48\sqrt{2}) \frac{\sqrt{m}}{\sqrt{D_{\mathrm{eff}}}}.$$

□

3.4 CORR-T3: high-probability supremum tail (Lean-exported form)

Theorem 3.5 (CORR-T3)**.** *Assume $m \geq 1$, $L \geq 1$, $\rho_{\mathrm{sum}} > 0$, and $t > 0$. Then*

$$\mathbb{P}\left(t \leq \sup_{u \in S^{m-1}} X_{\mathrm{corr}}(u)\right) \leq 2m \exp\left(-\frac{(t/m)^2 D_{\mathrm{eff}}}{2}\right).$$

Proof. Write $X_{\mathrm{corr}}(u) = \langle u, W_{\mathrm{corr}} \rangle$. If $|W_{\mathrm{corr},i}| < t/m$ for all i, then

$$\sup_{u \in S^{m-1}} X_{\mathrm{corr}}(u) \leq \sum_{i=1}^{m} |W_{\mathrm{corr},i}| < m \cdot (t/m) = t,$$

so

$$\left\{t \le \sup_u X_{\mathrm{corr}}(u)\right\} \subseteq \bigcup_{i=1}^{m} \{|W_{\mathrm{corr},i}| \ge t/m\}.$$

For each coordinate basis vector e_i, Lean proves $W_{\mathrm{corr},i} = M^{\mathrm{flat}}_{e_i}$. By Theorem **??** with $u = e_i$,

$$\mathbb{P}\left(|W_{\mathrm{corr},i}| \ge t/m\right) \le 2\exp\left(-\frac{(t/m)^2 D_{\mathrm{eff}}}{2}\right).$$

Union bound over $i = 1, \dots, m$ gives the result. □

4 Relation to Recent ArXiv Work

4.1 Inverse-depth scaling and effective depth

The IID model yields $\mathrm{Var}(M_u) = 1/L^2$ for $\|u\| = 1$, reflecting both the ResNet scaling $1/\sqrt{L}$ in each layer increment and the explicit depth average $1/L$. The correlated model replaces L^2 by the effective depth $D_{\mathrm{eff}} = L^3/\rho_{\mathrm{sum}}$, which interpolates between $D_{\mathrm{eff}} = L^2$ (independent layers) and $D_{\mathrm{eff}} = L$ (perfectly correlated layers). This formalizes, in a Gaussian-process setting, the same qualitative mechanism proposed in [**?**]: when most layers are functionally similar, depth behaves like averaging over fewer effectively independent contributions, producing inverse-depth-type scaling governed by an effective depth rather than the nominal depth.

4.2 Lean formalization of empirical-process tools

The sphere-process bounds in Theorems **??** and **??** are applications of Dudley's entropy integral theorem for sub-Gaussian processes. The explicit Dudley constant $12\sqrt{2}$ and the entropy-integral interface used here align with the Lean 4 empirical-process infrastructure developed in [**?**], which formalizes Dudley's theorem for sub-Gaussian processes (and related Gaussian concentration tools) in Lean. The present note specializes these tools to the NNGPResNet sphere process and records a one-to-one correspondence with the concrete Lean declarations used in our development.

5 Lean Correspondence

Every mathematical object/theorem used above is matched below to the exact Lean declaration and file.

IID model definitions and bridges

- Residual width $rm = rm$: `LeanSlt.NNGPResNet.residualWidth` — `LeanSlt/NNGPResNet/Defs.lean`
- Sign activation sign: `LeanSlt.NNGPResNet.signAct` — `LeanSlt/NNGPResNet/Defs.lean`
- Layer increment Δh_ℓ: `LeanSlt.NNGPResNet.DeltaH` — `LeanSlt/NNGPResNet/Defs.lean`
- Layer probe $g_\ell(u)$: `LeanSlt.NNGPResNet.g` — `LeanSlt/NNGPResNet/Defs.lean`
- Depth average M_u: `LeanSlt.NNGPResNet.M_u` — `LeanSlt/NNGPResNet/Defs.lean`
- Averaged vector W: `LeanSlt.NNGPResNet.W` — `LeanSlt/NNGPResNet/Defs.lean`

- Sphere type S^{m-1}: `LeanSlt.NNGPResNet.Sphere` — `LeanSlt/NNGPResNet/Defs.lean`
- Process $X(u) = \langle u, W \rangle$: `LeanSlt.NNGPResNet.X` — `LeanSlt/NNGPResNet/Defs.lean`
- Bridge $M_u = X(u)$: `LeanSlt.NNGPResNet.M_u_eq_X` — `LeanSlt/NNGPResNet/Defs.lean`
- Bridge $M_u = \langle u, W \rangle$: `LeanSlt.NNGPResNet.M_u_eq_inner_W` — `LeanSlt/NNGPResNet/Defs.lean`

IID scalar results (T1)

- Gaussian law: `LeanSlt.NNGPResNet.scalar_mean_map_gaussian` — `LeanSlt/NNGPResNet/Scalar.lean`
- Tail bound: `LeanSlt.NNGPResNet.scalar_mean_tail` — `LeanSlt/NNGPResNet/Scalar.lean`
- Second moment: `LeanSlt.NNGPResNet.scalar_mean_mse` — `LeanSlt/NNGPResNet/Scalar.lean`

IID sphere process, entropy, Dudley, high probability (T2/T3)

- Sphere set and subtype bridge: `sphereSet`, `SphereSetSub`, `sphereEquivSphereSetSub` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Increment Gaussian law: `LeanSlt.NNGPResNet.map_X_diff_gaussian` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Sub-Gaussian process instance: `LeanSlt.NNGPResNet.sphere_process_isSubGaussian` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Sphere entropy finiteness: `LeanSlt.NNGPResNet.entropyIntegralENNReal_sphere_ne_top` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Sphere entropy bound $\leq 4\sqrt{m}$: `LeanSlt.NNGPResNet.entropyIntegral_sphere_le` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Dudley intermediate bound: `LeanSlt.NNGPResNet.sphere_process_expected_sup_bound_entropy` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Entropy-assuming intermediate theorem: `LeanSlt.NNGPResNet.sphere_process_expected_sup_bound_assuming_entropy` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Final unconditional IID-T2 theorem: `LeanSlt.NNGPResNet.sphere_process_expected_sup_bound` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Coordinate bridge: `LeanSlt.NNGPResNet.W_coord_eq_M_u_single` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Supremum event reduction to coordinates: `LeanSlt.NNGPResNet.sphere_sup_event_subset_union_coords` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Integrability and positivity helpers used in T3 proof path: `LeanSlt.NNGPResNet.integrable_supX`, `LeanSlt.NNGPResNet.supX_nonneg` — `LeanSlt/NNGPResNet/SphereChaining.lean`
- Final IID-T3 theorem: `LeanSlt.NNGPResNet.sphere_process_highProb_bound` — `LeanSlt/NNGPResNet/SphereChaining.lean`

Correlated definitions and effective depth

- Correlation kernel ρ: `LeanSlt.NNGPResNet.Correlated.rho` — `LeanSlt/NNGPResNet/Correlated/Defs.lean`
- Correlation sum ρ_{sum}: `LeanSlt.NNGPResNet.Correlated.rhoSum` — `LeanSlt/NNGPResNet/Correlated/Defs.lean`
- Inverse effective depth ρ_{sum}/L^3: `LeanSlt.NNGPResNet.Correlated.invDEff` — `LeanSlt/NNGPResNet/Correlated/Defs.lean`
- Effective depth L^3/ρ_{sum}: `LeanSlt.NNGPResNet.Correlated.D_eff` — `LeanSlt/NNGPResNet/Correlated/Defs.lean`
- Layer probe g_ℓ^{corr}: `LeanSlt.NNGPResNet.Correlated.gCorr` — `LeanSlt/NNGPResNet/Correlated/Defs.lean`
- Layer-average statistic: `LeanSlt.NNGPResNet.Correlated.M_u_corr_layerAvg` — `LeanSlt/NNGPResNet/Correlated/Defs.lean`
- Flattened statistic: `LeanSlt.NNGPResNet.Correlated.M_u_corr` — `LeanSlt/NNGPResNet/Correlated/Defs.lean`
- Bridge lemma: `LeanSlt.NNGPResNet.Correlated.M_u_corr_layerAvg_eq_M_u_corr` — `LeanSlt/NNGPResNet/Correlated/Defs.lean`
- Correlated averaged vector and process: `LeanSlt.NNGPResNet.Correlated.Wcorr`, `LeanSlt.NNGPResNet.Correlated.XCorr` — `LeanSlt/NNGPResNet/Correlated/Defs.lean`

Correlated covariance/variance/scalar results

- Single-layer Gaussian map: `LeanSlt.NNGPResNet.Correlated.map_gCorr_gaussian_any` — `LeanSlt/NNGPResNet/Correlated/GaussianLinearForms.lean`
- Single-layer variance: `LeanSlt.NNGPResNet.Correlated.variance_gCorr` — `LeanSlt/NNGPResNet/Correlated/GaussianLinearForms.lean`
- Variance of layer sum: `LeanSlt.NNGPResNet.Correlated.variance_gCorr_add` — `LeanSlt/NNGPResNet/Correlated/GaussianLinearForms.lean`
- Covariance formula (CORR-COV): `LeanSlt.NNGPResNet.Correlated.cov_gCorr` — `LeanSlt/NNGPResNet/Correlated/GaussianLinearForms.lean`
- Unit-vector Gaussian law (layerAvg): `LeanSlt.NNGPResNet.Correlated.corr_scalar_map_gaussian_unit_layerAvg` — `LeanSlt/NNGPResNet/Correlated/Scalar.lean`
- Unit-vector tail (layerAvg): `LeanSlt.NNGPResNet.Correlated.corr_scalar_tail_layerAvg` — `LeanSlt/NNGPResNet/Correlated/Scalar.lean`
- Unit-vector MSE (layerAvg): `LeanSlt.NNGPResNet.Correlated.corr_scalar_mse_layerAvg` — `LeanSlt/NNGPResNet/Correlated/Scalar.lean`
- Variance from covariance sum: `LeanSlt.NNGPResNet.Correlated.var_M_u_corr_layerAvg_as_covariance_sum` — `LeanSlt/NNGPResNet/Correlated/Scalar.lean`

- Variance in explicit ρ-sum form: `LeanSlt.NNGPResNet.Correlated.var_M_u_corr_layerAvg_as_rho_sum` — `LeanSlt/NNGPResNet/Correlated/Scalar.lean`

Correlated chaining and high-probability supremum

- Centered correlated process: `LeanSlt.NNGPResNet.Correlated.centeredCorrProcess` — `LeanSlt/NNGPResNet/Correlated/SphereChaining.lean`
- Correlated Dudley entropy form: `LeanSlt.NNGPResNet.Correlated.corr_sphere_expected_sup_bound_entropy` — `LeanSlt/NNGPResNet/Correlated/SphereChaining.lean`
- Correlated expected sup bound (flat form): `LeanSlt.NNGPResNet.Correlated.corr_sphere_expected_sup_bound` — `LeanSlt/NNGPResNet/Correlated/SphereChaining.lean`
- Correlated expected sup bound (layerAvg form): `LeanSlt.NNGPResNet.Correlated.corr_sphere_expected_sup_bound_layerAvg` — `LeanSlt/NNGPResNet/Correlated/SphereChaining.lean`
- Coordinate bridge for correlated process: `LeanSlt.NNGPResNet.Correlated.Wcorr_coord_eq_M_u_corr_single` — `LeanSlt/NNGPResNet/Correlated/SphereChaining.lean`
- Final correlated high-probability theorem: `LeanSlt.NNGPResNet.Correlated.corr_sphere_highProb_bound` — `LeanSlt/NNGPResNet/Correlated/SphereChaining.lean`

Module aggregators

- IID exports: `LeanSlt.NNGPResNet.Main` — `LeanSlt/NNGPResNet/Main.lean`
- Correlated exports: `LeanSlt.NNGPResNet.Correlated.Main` — `LeanSlt/NNGPResNet/Correlated/Main.lean`
- Top-level import hook: `LeanSlt.lean`.

References

[1] Y. Liu, S. Kangaslahti, Z. Liu, and J. Gore. Inverse depth scaling from most layers being similar. *arXiv preprint arXiv:2602.05970*, 2026.

[2] Y. Zhang, J. D. Lee, and F. Liu. Statistical learning theory in Lean 4: empirical processes from scratch. *arXiv preprint arXiv:2602.02285*, 2026.