Spectral Estimation with Deformed Decompression

Siavash Ameli [1,2] Chris van der Heide [3] Liam Hodgkinson [4]
Michael W. Mahoney [1,2,5]

[1] *Department of Statistics, UC Berkeley*
[2] *International Computer Science Institute*
[3] *Dept. of Electrical and Electronic Eng., University of Melbourne*
[4] *School of Mathematics and Statistics, University of Melbourne*
[5] *Lawrence Berkeley National Laboratory*

NeurIPS 2026

Motivation and Objective

Motivation:

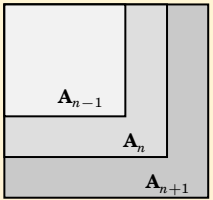
Spectra of (**generalized-**) **covariance** matrices reveal:

- Conditioning and near-zero directions.
- Spectral densities, edges, and gaps.
- Log-determinants and inverse traces.

Challenges:

- Direct eigendecomposition costs $\mathcal{O}(n^3)$.
- Matrix formation hits the $\mathcal{O}(n^2)$ memory wall.

Spectral Decompression



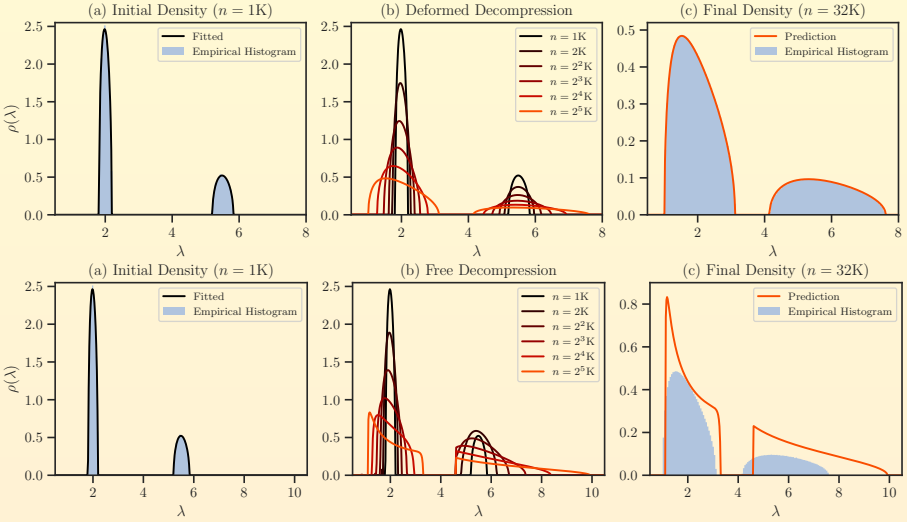
Consider a sequence of nested covariance matrices

$$\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots, \qquad \mathbf{A}_n \in \mathbb{R}^{n \times n}.$$

Objective: Infer the spectral distribution of a **large covariance matrix** $\mathbf{A}_n$ from a much smaller, **representative principal submatrix** $\mathbf{A}_{n_0}$, where $n_0 \ll n$.

Key Question: Can we predict how the spectrum evolves with matrix size without knowing the population covariance?

Why "Deformed" Decompression?



- *Top:* **Deformed Decompression (DD)** — ***this work***: correctly predicts the large-matrix spectrum.
- *Bottom:* **Free Decompression (FD)** — Ameli et al., NeurIPS 2025 & 2026: fails for this covariance spectrum.

Covariance Family:

$$\mathbf{S}_p = \frac{1}{n}\mathbf{X}\mathbf{X}^\intercal, \qquad \mathbf{X} = \mathbf{\Sigma}_p^{1/2}\mathbf{Z}, \qquad c = \frac{p}{n}.$$

- Increase the aspect ratio $c \to \tau c$, preserving the limiting population spectral law.
- The population covariance $\mathbf{\Sigma}_p$ is **unknown**.

Stieltjes and Companion Transforms:

$$m(z) = \lim_{p\to\infty} \frac{1}{p}\,\mathrm{trace}(\mathbf{S}_p - z\mathbf{I})^{-1}, \qquad \bar{m}(z) = c\,m(z) + \frac{c-1}{z}.$$

R-transform: Defined by $R(-m(z)) = z + m(z)^{-1}$, likewise for $\bar{R}$.

Our Theorem (Deformed Decompression)

For the **companion** R-transform:

$$\bar{R}_\tau(w) = \tau \bar{R}_1(w).$$

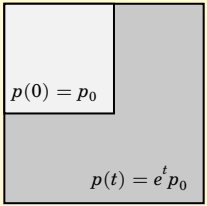
Nica–Speicher (Free Decompression)

For the **ordinary** R-transform:

$$R_\tau(w) = R_1(\tau w).$$

Proposition (Distinct Decompression Flows)

DD and FD can coincide only for scaled ***Marchenko–Pastur*** *laws.*



- Consider covariance matrices $\mathbf{S}_{p(t)}$ of dimension $p(t) = e^t p_0$ and fixed sample size n.
- The aspect ratio evolves as $c(t) = e^t c_0$.
- We compare two distinct decompression laws for different matrix families:

Our Theorem (Deformed Decompression)

Covariance matrices: **companion** Stieltjes transform

$$\frac{\partial \bar{m}}{\partial t} = -\left(z + \frac{1}{\bar{m}}\right) \frac{\partial \bar{m}}{\partial z}$$

Ameli et al. (NeurIPS 2025, Free Decompression)

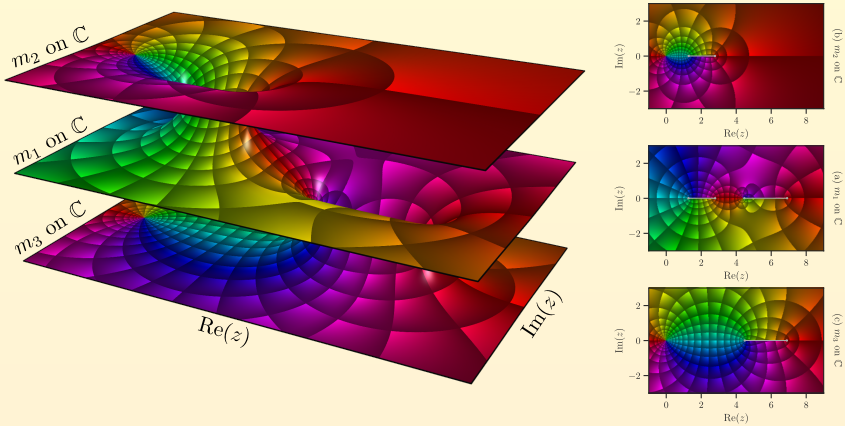
Asymptotically free matrices: Stieltjes transform

$$\frac{\partial m}{\partial t} = -m + \frac{1}{m} \frac{\partial m}{\partial z}$$

Further Results:

- **Stability:** Initial errors grow at most polynomially: $\|\eta_\tau\|_{L^2} \leq \tau^\nu \|\eta_1\|_{L^2}$.
- **Generalized Covariance Matrices:** Extends DD to PSD Hessian-type matrices, with approximation error $O(\sigma_D^2)$.

Algebraic Spectral Curve



- ***Algebraic Assumption*** (Ameli et al., NeurIPS 2026): The Stieltjes transform satisfies $P(z, m) = 0$.
- *Right:* Three algebraic branches of a cubic deformed Marchenko–Pastur law with two spectral bulks.
- *Left:* Gluing these branches across cuts forms a **spectral curve**: $\mathcal{C} = \{(z, m) \in \mathbb{C}^2 \mid P(z, m) = 0\}$
- ***Our Approach:*** Evolve the spectral curve under the new DD law.

Deformed Decompression on Algebraic Spectral Curves

Theorem (Algebraic Deformed Decompression)

Suppose $P(z,m)=0$. *Define the companion spectral curve*

$$\bar{P}(\zeta,w)=P\left(\zeta,c^{-1}\left(w-\zeta^{-1}(c-1)\right)\right).$$

Then, under DD,

$$\bar{P}(\zeta,\bar{m}_\tau(z))=0,\quad \zeta=\frac{z}{\tau}+\frac{\tau^{-1}-1}{\bar{m}_\tau(z)}.$$

Advantages

- Algebraic structure is **closed** under DD.
- Avoids PDE and ill-posed analytic continuation.
- Handles multi-bulk and atomic covariance spectra.

Numerical Realization

- Reuse polynomial fitting of Ameli et al. (NeurIPS 2026).
- Track the physical branch along the *new* DD flow.

Beyond the Density

Direct Evolution of Spectral Quantities: without reconstructing the full density. Define

$$F_1=\bar{P}(\zeta,w),\qquad F_2=\tau w^2\partial_w\bar{P}+(\tau-1)\partial_\zeta\bar{P}.$$

Spectral Edges

$$F_1=F_2=0$$

Track spectral bulk boundaries.

Cusps & Bulk Splitting

$$F_1=F_2=0$$
$$\det\partial(F_1,F_2)/\partial(\zeta,w)=0$$

Identify splitting time/location.

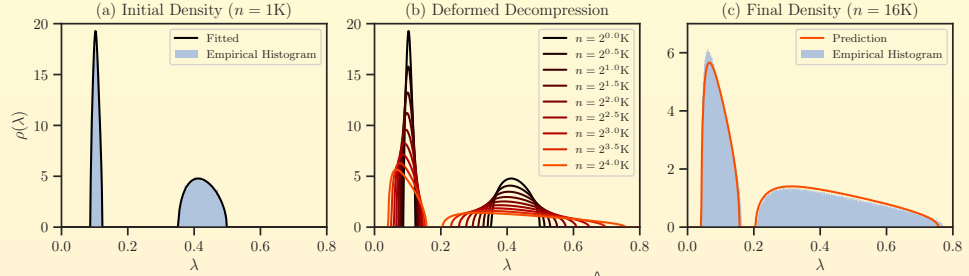
Nonzero Spectral Atoms

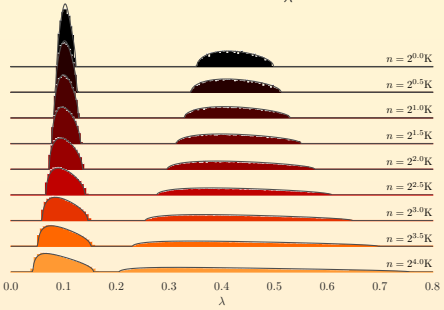
$$x_\circ(\tau)=\tau x_\circ$$
$$w(\tau)=c^{-1}+\tau(w_\circ-c^{-1})$$

Track atom locations and masses.

Spectral Moments: Direct evolution from initial moments, without reconstructing the density.

Examples: Hessian I

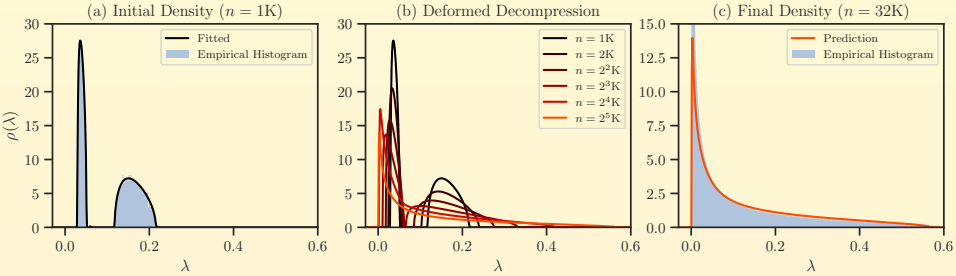




Logistic-Regression Hessian

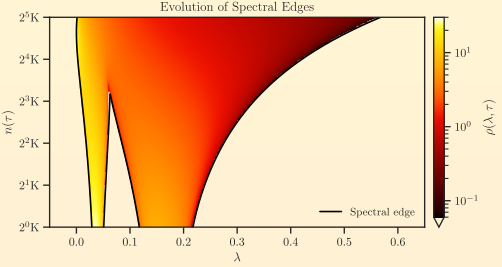
- **120K samples**, 16K-dimensional Hessian.
- DD extrapolates from **1K to 16K**.
- **Right:** Ridge plot of density evolution.

Examples: Hessian II

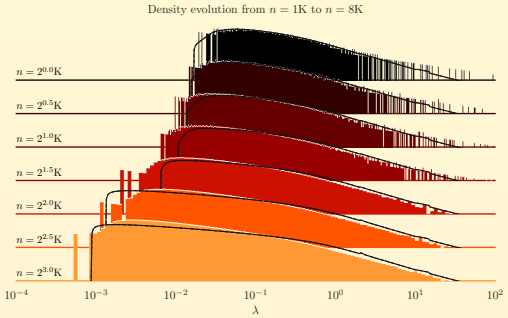


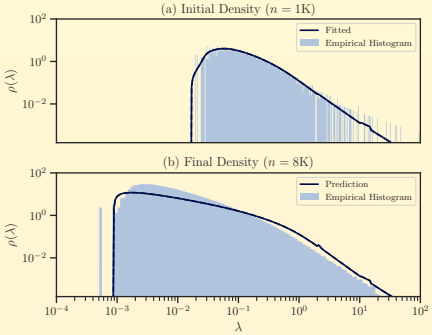
Logistic-Regression Hessian

- **100K samples**, 32K-dimensional Hessian.
- DD extrapolates from **1K to 32K**.
- **Right:** Direct evolution of spectral edges captures the **merging** of two bulks.



Examples: STL Dataset





- ***STL-10 Dataset:*** Gram covariance matrix of 8K images with 27,648 features.
- DD extrapolates from **1K to 8K**.
- Density evolution (left), initial and final spectra (right).

Resources

This Work

Spectral Estimation with Deformed Decompression
Ameli, S., van der Heide, C., Hodgkinson, L., Mahoney, M.W. *NeurIPS 2026*

Related Works

Free Decompression with Algebraic Spectral Curves
Ameli, S., van der Heide, C., Hodgkinson, L., Mahoney, M.W. *NeurIPS 2026*

Spectral Estimation with Free Decompression
Ameli, S., van der Heide, C., Hodgkinson, L., Mahoney, M.W. *NeurIPS 2025*

Software

Package	Documentation	Install
freealg	`ameli.github.io/freealg`	`pip install freealg`