Learning Geometric Residual Structure in Galaxy Rotation Curves

Jonathan R. Landers
0000-0003-1872-6179
Independent Researcher

Abstract

Galaxy rotation curves define a noisy inverse problem: the baryonic mass model predicts one acceleration field, while the observed dynamics imply another. This paper asks whether the leftover field after fitting a physical baseline is compressible across galaxies. We introduce a masked, smooth, nonnegative dictionary-learning diagnostic for SPARC rotation-curve residuals. For each galaxy we compute the baryon-subtracted acceleration residual and the residual-of-residual left after fitting NFW to the outer region; a two-channel lift represents signed residual fields with nonnegative coefficients. A theorem shows that, under an idealized baseline-projection model, residual-of-residual learning removes arbitrary baseline variation while preserving latent misspecification modes. On 131 primary-quality SPARC galaxies, a rank-six dictionary reduces held-out radial profile error from 0.543 to 0.098 for positive acceleration residuals and from 1.075 to 0.168 for signed NFW residual-of-residuals. The first three signed modes are bootstrap-stable, and learned coefficients align with an independent central-overshoot diagnostic. These results do not decide between dark matter and modified gravity. They show that population residual modes are learnable model-criticism objects for the missing-mass inverse problem.

1. Introduction

Flat galaxy rotation curves have long indicated that the geometry inferred from dynamics is not the geometry predicted from observed baryons alone [1, 2]. The usual next step is to choose a physical source family and fit it. The dominant explanation is a non-baryonic dark component whose halos are often modeled with the NFW profile motivated by cold-dark-matter simulations [3, 4]. However, disk-galaxy mass modeling also exhibits persistent small-scale structure: cored profiles often fit individual galaxies well [5, 6], observed dwarf rotation curves are more diverse than many baseline simulations predict [7, 8], and the empirical coupling between baryonic and dynamical acceleration is striking [9, 10]. Recent weak-lensing analyses have further extended the empirical reach of flat circular velocities [11]. These facts do not imply that a simple alternative to cold dark matter is correct. They do imply that repeated rotation-curve residuals should be treated as structured data, not merely nuisance errors.

This paper treats the problem as a physics-constrained inverse problem [12]. Let

$$\Delta g(R) = \frac{V_{\rm obs}^2(R) - V_{\rm bar}^2(R)}{R} \tag{1}$$

be the weak-field acceleration residual. A named source model supplies a candidate residual $\Delta g_\theta(R)$; the object studied here is the leftover field

$$\epsilon_g(R) = \Delta g(R) - \Delta g_\theta(R). \tag{2}$$

This is a residual-of-residual: the part of the inferred geometry not accounted for by the chosen baseline source. The logic is conservative: fit a physical baseline, remove it, and learn the recurring structure that remains. In that sense ϵ_g is a model-criticism object [13], not a direct theory selector.

The distinction matters empirically in the SPARC sample studied below. Before learning, the same residual pipeline finds that cored/isothermal residual profiles beat NFW for 80.2% of primary-quality galaxies, and outer-NFW fits leave central overshoot in 51.9%. These quantities are not training labels. They are held back as independent diagnostics for asking whether learned population coordinates recover recognizable residual structure.

The geometric point is that Δg is not just a scalar error bar or a goodness-of-fit residual. In the weak-field limit it is the radial derivative of an inferred potential perturbation, $\Delta g = \mathrm{d}\delta\Phi/\mathrm{d}R$. Thus a coherent pattern in ϵ_g is a coherent leftover field after a named source model has already been given the opportunity to explain the outer rotation curve. The learning problem is to ask whether those leftover fields share a low-dimensional population structure before assigning them a physical cause.

Proposition 1 (Flat-tail residual geometry)**.** *On a radial interval where the excess circular speed is approximately flat,*

$$V_{\rm obs}^2(R) - V_{\rm bar}^2(R) = v_f^2, \tag{3}$$

the residual acceleration, weak-field potential perturbation, and spherical effective-density proxy are

$$\Delta g(R) = \frac{v_f^2}{R}, \tag{4}$$

$$\delta\Phi(R) = v_f^2 \log(R/R_0) + C, \tag{5}$$

$$\rho_{\rm eff}(R) = \frac{v_f^2}{4\pi G R^2}. \tag{6}$$

Proof. The circular-speed relation gives

$$\Delta g(R) = \frac{V_{\rm obs}^2(R) - V_{\rm bar}^2(R)}{R} = \frac{v_f^2}{R}. \tag{7}$$

Because

$$\Delta g(R) = \frac{\mathrm{d}\delta\Phi}{\mathrm{d}R}, \tag{8}$$

integration gives the logarithmic potential. For the spherical proxy,

$$M_{\mathrm{eff}}(< R) = \frac{R^2 \Delta g(R)}{G} = \frac{v_f^2 R}{G}, \tag{9}$$

and therefore

$$4\pi R^2 \rho_{\mathrm{eff}}(R) = \frac{\mathrm{d}M_{\mathrm{eff}}}{\mathrm{d}R} = \frac{v_f^2}{G}. \tag{10}$$

□

This elementary proposition is useful because it explains why outer rotation curves are a hard baseline test. A profile family can match a flat tail by approximating a logarithmic residual over a finite radial band, yet still imply the wrong inner residual when extrapolated inward. The experiment below uses that failure mode as a measurable diagnostic rather than as a final physical verdict.

The machine-learning analogy is source disaggregation. In non-intrusive load monitoring, an aggregate power trace is decomposed into latent appliance contributions [14, 15]. Here the aggregate is an irregular radial field, and the latent components are geometric mismatch modes. Because the profiles are sparse, uncertain, signed, and expected to be smooth, we use dictionary learning and nonnegative matrix factorization (NMF) [16–18], not a temporal appliance model.

Contributions:

1. We formulate population residual learning for galaxy rotation curves as a masked, smooth, nonnegative dictionary problem on a normalized radial grid.
2. We introduce a signed two-channel lift for residual-of-residual fields.
3. We prove a baseline-projection theorem explaining why learning residual-of-residuals can isolate misspecification modes hidden by baseline variation.
4. We run the method on public SPARC residuals [19], compare learned modes with named templates, and bootstrap mode stability.

2. Closest Literature

The closest astrophysical literature is the part of halo fitting concerned with repeated residual morphology. NFW halos provide a simulation-motivated baseline [3, 4], while cored profiles and observed dwarf-galaxy diversity expose coherent departures from simple cuspy fits [5–7]. The radial acceleration relation and MOND phenomenology emphasize the empirical coupling between baryonic and dynamical accelerations [9, 10, 20]. Our experiment does not replace these theories with a learned black box. It uses learning to summarize the field that remains after one named baseline has been fit.

The closest statistical literature is inverse-problem model criticism. Rotation curves do not directly observe a source density; they observe a dynamical field from which source models are inferred under assumptions about geometry, stellar mass-to-light ratios, distance, inclination, and priors. Inverse-problem theory separates the forward operator from the data-misfit object [12]. Posterior predictive checking similarly treats structured discrepancies as scientific evidence about model inadequacy [13]. The residual-of-residual ϵ_g is a frequentist, profile-level discrepancy designed for this purpose.

The closest machine-learning literature is additive representation learning. NILM decomposes an aggregate power trace into appliance-level components [14, 15]. Sparse coding and dictionary learning decompose signals into reusable atoms [17, 18, 21]; NMF adds nonnegativity and parts-based interpretability [16]. The difference here is the observation geometry: the "aggregate" is an irregular radial residual profile, and the components are population-level mismatch modes rather than appliances.

3. Residual Learning Setup

3.1. Residual profiles

SPARC provides rotation curves and baryonic velocity components for 175 disk galaxies [19]. We use fixed stellar mass-to-light ratios $\Upsilon_{\mathrm{disk}} = 0.5$ and $\Upsilon_{\mathrm{bulge}} = 0.7$, following common SPARC exploratory practice, and compute

$$V_{\mathrm{bar}}^2 = \mathrm{sign}(V_{\mathrm{gas}}) V_{\mathrm{gas}}^2 + \Upsilon_{\mathrm{disk}} V_{\mathrm{disk}}^2 + \Upsilon_{\mathrm{bulge}} V_{\mathrm{bulge}}^2. \tag{11}$$

For each usable galaxy we fit NFW, cored/isothermal, and logarithmic-tail residual families. The dictionary experiment uses the NFW fit restricted to the outer radial region as the baseline. This is the central stress test: if an NFW source is forced to match the flat outer residual, does it leave coherent inner structure behind? This complements Bayesian SPARC halo catalogs that fit complete halo families under priors [22].

Profiles are placed on a common grid $x = R/R_d$, where R_d is the SPARC disk scale length. Each galaxy is normalized by the root-mean-square of its observed residual so that the dictionary learns shape rather than halo scale. Let $X \in \mathbb{R}^{n\times m}$ denote the nonnegative matrix for $\max(\Delta g, 0)$ and let $E \in \mathbb{R}^{n\times m}$ denote the signed residual-of-residual matrix. Missing entries arise because galaxies cover different radial ranges; a mask M identifies observed grid entries.

3.2. Dictionary objectives

For positive residuals we fit

$$\min_{A,\Phi\geq 0} \sum_{ij} M_{ij} w_{ij} \left(X_{ij} - (A\Phi)_{ij}\right)^2 + \lambda \sum_k \|\nabla_x^2 \phi_k\|_2^2, \tag{12}$$

with $A \in \mathbb{R}_+^{n\times K}$ and $\Phi \in \mathbb{R}_+^{K\times m}$. For signed residuals, we use the two-channel lift

$$E = E^+ - E^-, \qquad E^\pm \geq 0, \tag{13}$$

and learn a nonnegative dictionary on $[E^+, E^-]$. The decoded signed mode is $\phi_k = \phi_k^+ - \phi_k^-$. This preserves

additive, nonnegative galaxy coefficients while allowing each learned mode to have positive and negative radial structure.

Ranks $K = 1, \ldots, 6$ are scored by holding out 15% of observed radial grid entries per split. Rank zero is the population-mean baseline. Bootstrap stability is measured by resampling galaxies, refitting the selected dictionary, and matching learned modes to the reference modes by absolute profile correlation.

3.3. Optimization protocol

The masked objective is optimized by multiplicative NMF-style updates with small denominators clipped for numerical stability. With Hadamard product $\odot$ and weight matrix W, the positive-residual update has the form

$$A \leftarrow A \odot \frac{(M \odot W \odot X)\Phi^\top}{(M \odot W \odot A\Phi)\Phi^\top + \varepsilon}, \tag{14}$$

$$\Phi \leftarrow \Phi \odot \frac{A^\top(M \odot W \odot X)}{A^\top(M \odot W \odot A\Phi) + \varepsilon}. \tag{15}$$

The signed experiment applies the same updates to the lifted matrix $[E^+, E^-]$. After each fit, mode profiles are smoothed by a second-difference penalty, normalized, and paired with rescaled galaxy coefficients. Each rank is fit from multiple random initializations; the seed with lowest masked training loss is used for held-out scoring. Bootstrap refits repeat the full selection and mode-matching procedure, so stability reflects both sample variation and optimizer variation.

4. Why Learn Residual-of-Residuals?

The following idealization isolates the statistical reason for subtracting the baseline before learning. If baseline variation is large, raw residual learning can spend its capacity on the fitted family itself rather than on the structure that the family missed.

Theorem 1 (Baseline-projected residual mode identifiability)**.** *Let $\mathcal{H}$ be a finite-dimensional residual-profile space, let $\mathcal{B} \subset \mathcal{H}$ be a linear baseline subspace, and let $\Pi_{\mathcal{B}}$ be the orthogonal projection. For each galaxy, write*

$$x_i = b_i + s_i + \eta_i, \qquad b_i \in \mathcal{B}, \quad s_i \in \mathcal{B}^\perp, \tag{16}$$

where b_i is baseline variation, s_i is latent misspecification, and η_i is noise. Define

$$e_i = (I - \Pi_{\mathcal{B}})x_i. \tag{17}$$

Then:

1. Projection removes the baseline:

$$e_i = s_i + (I - \Pi_{\mathcal{B}})\eta_i. \tag{18}$$

2. If $\eta_i = 0$ and the misspecification has a conic dictionary

$$s_i = \sum_{k=1}^{K} a_{ik}\phi_k, \qquad a_{ik} \geq 0, \quad \phi_k \in \mathcal{B}^\perp, \tag{19}$$

then the residual-of-residual has the same dictionary:

$$e_i = \sum_{k=1}^{K} a_{ik}\phi_k. \tag{20}$$

3. In centered noiseless data, with row matrices $X = B + S$,

$$E = (I - \Pi_{\mathcal{B}})X = S, \qquad \operatorname{range}(E^\top E) = \operatorname{span}\{s_i\}. \tag{21}$$

The raw covariance after scaling $b_i \mapsto cb_i$ is

$$C_X(c) = c^2 B^\top B + S^\top S. \tag{22}$$

If

$$\alpha = \lambda^+_{\min}(B^\top B), \qquad \beta = \lambda_{\max}(S^\top S), \tag{23}$$

then the threshold

$$c^2\alpha > \beta \tag{24}$$

makes every nonzero baseline eigenvalue exceed every misspecification eigenvalue. As $c \to \infty$, the leading raw eigenspace converges to the baseline data eigenspace inside $\mathcal{B}$.

Proof. Since $\Pi_{\mathcal{B}} b_i = b_i$ and $\Pi_{\mathcal{B}} s_i = 0$,

$$(I - \Pi_{\mathcal{B}})x_i = (I - \Pi_{\mathcal{B}})(b_i + s_i + \eta_i) = s_i + (I - \Pi_{\mathcal{B}})\eta_i, \tag{25}$$

which proves the projection claim. The dictionary claim follows by substituting the noiseless conic expansion for s_i.

For the covariance claim, center the noiseless matrices and write $X = B + S$ with row spaces in $\mathcal{B}$ and $\mathcal{B}^\perp$. Orthogonality gives

$$B^\top S = S^\top B = 0, \tag{26}$$

so projection gives $E = S$ and

$$X(c)^\top X(c) = c^2 B^\top B + S^\top S. \tag{27}$$

This operator is block diagonal with respect to $\mathcal{B} \oplus \mathcal{B}^\perp$, so its nonzero eigenvalues are the union of the baseline and misspecification block eigenvalues. The condition $c^2\alpha > \beta$ orders all nonzero baseline directions ahead of every misspecification direction. Letting $c \to \infty$ leaves the leading raw eigenspace in $\mathcal{B}$, whereas projected residual learning has already removed $\mathcal{B}$. □

Remark 1. *NFW fitting is nonlinear and constrained, not an orthogonal projection. The theorem is not a claim of exact astrophysical identifiability. It is a model criticism principle: remove the fitted baseline first, then learn the coherent structure that remains.*

Table 1: Dataset summary for the dictionary experiment.

Quantity	Value
SPARC rotation-curve points	3391
Public SPARC galaxies	175
Dictionary-quality galaxies	131
Normalized radial grid points	72
Mean observed grid coverage	0.774
Central-overshoot fraction	0.519
Cored/isothermal beats NFW	0.802

Table 2: Held-out radial-point reconstruction. Rank zero is the population mean; $K = 6$ is the largest searched rank and was selected by held-out error.

Target	Baseline	$K = 6$	Improvement
$\max(\Delta g, 0)$	0.543	0.098	5.5×
ϵ_g signed	1.075	0.168	6.4×

5. Experiment

5.1. Data and preprocessing

The experiment uses residual profiles generated by the SPARC residual-of-residual pipeline. The mass-model table contains 3391 rotation-curve points for 175 galaxies. We use the primary-quality subset: SPARC quality flag 1 or 2, at least eight points, and inclination above 30° when available. After requiring finite disk scale length and enough normalized-radius coverage, the dictionary dataset contains 131 galaxies.

The common grid has 72 log-spaced points over $0.25 \leq R/R_d \leq 8.0$, with mean per-galaxy coverage 0.774. The same sample has a central-overshoot fraction of 0.519 under the independent NFW residual-of-residual diagnostic and a cored/isothermal-beats-NFW fraction of 0.802. These values match the earlier residual diagnostic and provide labels for post hoc interpretation only; they are not used to train the dictionary. Figure 1 shows the independent inner-gap diagnostic that supplies the central-overshoot label used below.

5.2. Evaluation

For each rank, five random radial holdout splits are used. The score reported below is normalized profile RMSE on held-out entries after decoding signed two-channel predictions. We also report bootstrap mode stability: 40 bootstrap resamples of galaxies are refit, and modes are matched by absolute correlation. Template correlations compare learned modes with simple NFW-like, cored, central, outer-tail, and overshoot templates.

6. Results

Table 2 gives the main compression result. The learned dictionary substantially outperforms the population-mean baseline for both positive residual profiles and signed NFW residual-of-residual profiles. The model-selection curve in Fig. 2 decreases monotonically through $K = 6$, so the experiment demonstrates compressible population structure rather than a sharp intrinsic rank. The selected rank should therefore be read as a useful search endpoint, not as a claim that galaxy residuals have exactly six physical components.

Table 3: Mode-level interpretation for the signed residual-of-residual dictionary. Correlations are post hoc diagnostics, not training labels.

Mode	Main empirical association	Value
1	Negative inner overshoot template	0.869
2	Sign-reversed overshoot template	0.936
3	Negative inner overshoot template	0.877
4	Central-overshoot flag	0.413
4	Normalized inner gap	−0.592
6	Central-overshoot flag	−0.371

Table 4: Bootstrap stability of selected signed residual-of-residual modes. Values are median absolute correlations to matched reference modes, with 10th and 90th percentiles in parentheses.

Mode	Median	10th pct.	90th pct.
1	0.997	0.967	1.000
2	0.967	0.838	0.996
3	0.976	0.919	0.997
4	0.846	0.593	0.987
5	0.684	0.579	0.858
6	0.890	0.626	0.949

The learned modes in Fig. 3 make the compression result interpretable. The positive-residual dictionary contains central, broad-tail, and cored-like components. The signed residual-of-residual dictionary contains several negative-inner-overshoot modes: cases where the NFW source required by the outer residual overpredicts the inner residual. Template matching confirms this: signed modes 1 and 3 correlate with a negative-inner-overshoot template at 0.869 and 0.877, while mode 2 matches the opposite sign of the same template with absolute correlation 0.936.

The first three signed modes are highly stable under galaxy bootstrap (Table 4); later modes are useful for reconstruction but less robust. Figure 4 shows the same stability pattern as matched correlation distributions. This suggests a conservative interpretation: the population has at least a few stable residual-of-residual modes, while higher-rank components should be treated as descriptive rather than discovery-level evidence.

The last check is whether the learned coordinates relate to diagnostics that were not used in fitting. For the signed residual-of-residual dictionary, coefficient 4 has correlation 0.413 with the binary central-overshoot flag and correlation −0.592 with the continuous normalized inner-gap statistic. Coefficient 6 has the opposite association with the overshoot flag (correlation −0.371). These associations are not used in training; they indicate that the learned coordinates recover part of the same structured failure mode found by direct NFW residual analysis. Figure 5

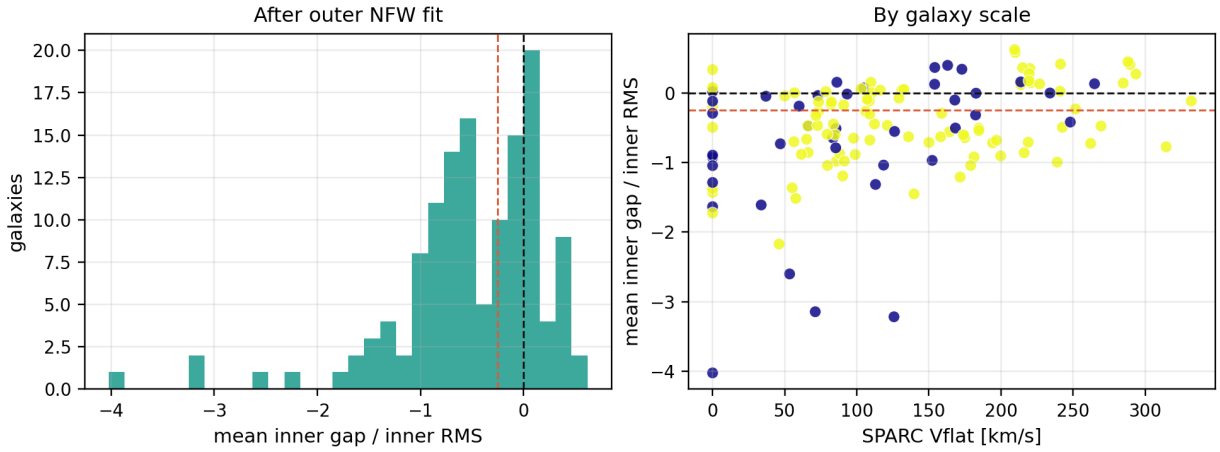

Figure 1: Independent NFW residual-of-residual diagnostic for the primary-quality SPARC sample. Left: distribution of the normalized mean inner gap after NFW has been fit to the outer region. Right: the same diagnostic versus SPARC flat velocity. The red dashed threshold is used only for post hoc interpretation of learned dictionary coefficients.

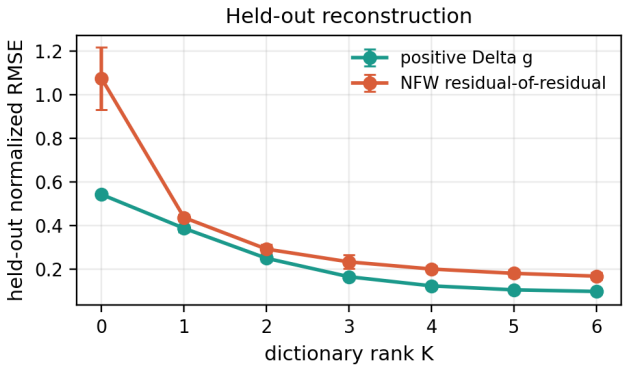

Figure 2: Held-out radial-point reconstruction as a function of dictionary rank. Both targets improve through the largest searched rank, indicating compressible residual structure without a clear elbow.

visualizes this learned coefficient space.

7. Discussion

The contribution is statistical model criticism. It does not show that NFW is wrong as a cosmological model, nor does it adjudicate between particle dark matter and modified dynamics. The experiment uses fixed mass-to-light ratios, simplified residual-space fits, no concentration prior, no baryonic feedback model, and no marginalization over distance or inclination. A full halo inference would require a richer likelihood and priors comparable to published SPARC halo catalogs [22].

The learning extension is therefore deliberately modest: learn the population structure of the discrepancy, not the final source model. This keeps the statistical task aligned with the inverse problem. If a mode is unstable, it is probably a nuisance or a sample artifact. If a mode is stable, appears in held-out radial entries, and correlates with independent residual diagnostics, then it becomes a concrete target for physical modeling. A simulator, field-equation surrogate, or baryonic-systematics model can be asked to generate that mode rather than to explain every galaxy from scratch.

This framing also gives clear failure conditions. The result would be weak if held-out reconstruction were close to the population mean, if bootstrap modes could not be matched across resamples, or if learned coefficients were orthogonal to all independent residual diagnostics. The present experiment passes those checks for the leading signed modes but not uniformly for every rank-six component. That is why the claim is deliberately about stable population residual structure, not about a discovered six-component law of galaxy dynamics.

The positive claim is that the residual left after fitting a physical baseline has learnable population structure. This is useful for three reasons. First, it provides a low-dimensional language for residual morphology, analogous to dictionary learning in signal processing [17, 21]. Second, it separates baseline fit quality from repeated misspecification patterns, in the spirit of posterior predictive checks and realized discrepancies [13]. Third, it supplies a bridge to more ambitious scientific machine-learning approaches: simulation-based inference could learn which physical generators produce the observed residual modes [23], while physics-informed models could add differentiable field-equation constraints [24].

The strongest near-term extension is held-out-galaxy validation. The present held-out radial-point score tests whether profiles are compressible after some entries are observed. A harder test is to predict the residual morphology of an entire unseen galaxy from baryonic covariates, environment, or partial radial coverage. Another necessary extension is uncertainty propagation over stellar mass-to-light ratios, distances, inclinations, and NFW concentration priors. If the stable signed modes persist under those perturbations, they become a specific target for physical explanation.

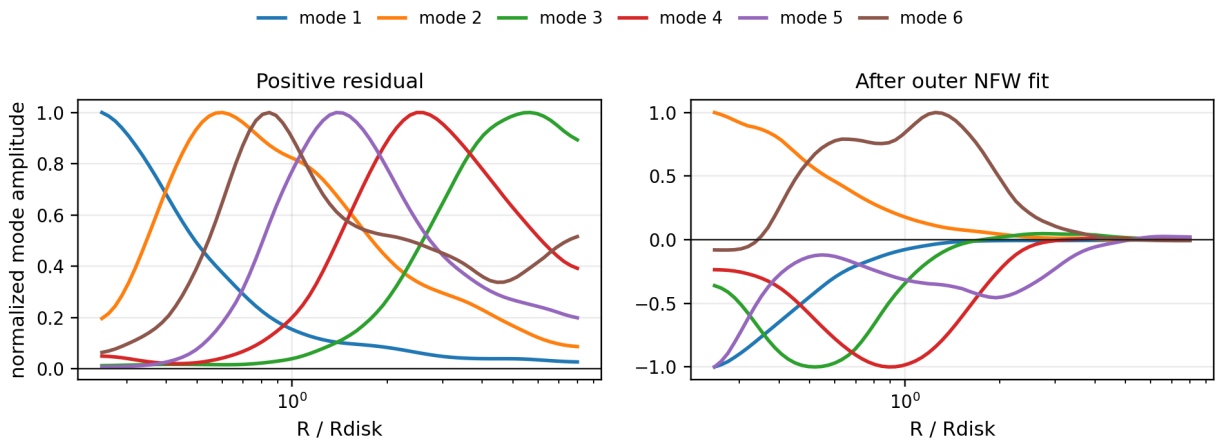

Figure 3: Learned population modes. Left: positive observed acceleration residual modes. Right: signed residual-of-residual modes after fitting NFW to the outer region. The signed modes are the more diagnostic object because they describe what remains after a physical baseline has been forced to match the outer residual.

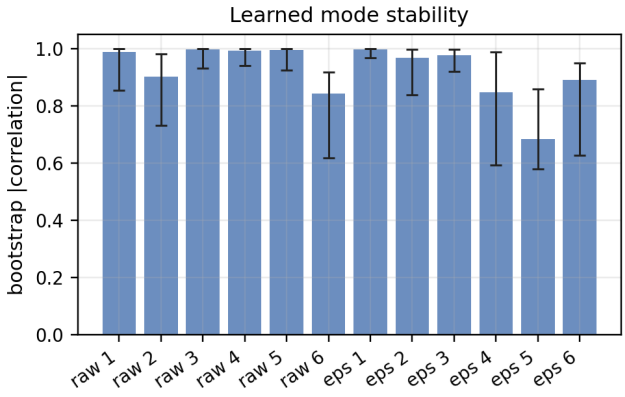

Figure 4: Bootstrap distribution of matched signed-mode correlations. The first three modes are stable across galaxy resamples; the later modes have larger sample sensitivity.

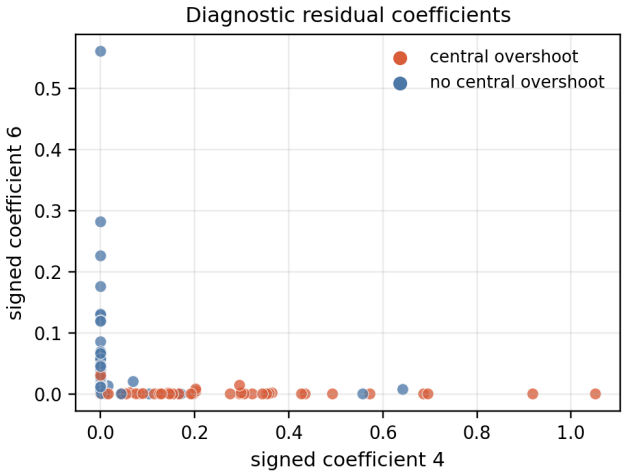

Figure 5: Signed residual-of-residual coefficient space. Points are galaxies; learned coordinates recover population variation that aligns with independent central-overshoot diagnostics even though those labels are withheld during training.

8. Conclusion

Learning geometric residuals across SPARC galaxies turns a familiar rotation-curve inverse problem into an interpretable statistical learning problem. A masked, smooth, nonnegative dictionary compresses both positive observed acceleration residuals and signed NFW residual-of-residuals. The signed dictionary is the central object: it learns coherent leftover fields after a baseline NFW source has been fit to the outer galaxy. The first signed modes are bootstrap-stable, and learned coefficients align with independent central-overshoot diagnostics. The resulting machine-learning framing is simple: fit the physical baseline, learn what remains, and use the learned residual geometry to guide the next physical model.

Reproducibility. The experiment is implemented in `experiments/sparc_residual_dictionary/`. It uses public SPARC tables and the residual profiles generated by the companion SPARC residual-of-residual pipeline, then writes all figures, CSV outputs, and the Markdown experiment report used by this paper.

References

[1] Vera C. Rubin, Jr. Ford, W. Kent, and Norbert Thonnard. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4 kpc) to UGC 2885 (R=122 kpc). *Astrophysical Journal*, 238: 471–487, 1980. doi: 10.1086/158003.

[2] Gianfranco Bertone and Dan Hooper. History of dark matter. *Reviews of Modern Physics*, 90:045002, 2018. doi: 10.1103/RevModPhys.90.045002.

[3] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The structure of cold dark matter halos. *Astrophysical Journal*, 462:563–575, 1996. doi: 10.1086/177173.

[4] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. A universal density profile from hierarchical clustering. *Astrophysical Journal*, 490:493–508, 1997. doi: 10.1086/304888.

[5] Andreas Burkert. The structure of dark matter halos in dwarf galaxies. *Astrophysical Journal Letters*, 447: L25–L28, 1995. doi: 10.1086/309560.

[6] W. J. G. de Blok. The core-cusp problem. *Advances in Astronomy*, 2010:789293, 2010. doi: 10.1155/2010/789293.

[7] Kyle A. Oman, Julio F. Navarro, Azadeh Fattahi, Carlos S. Frenk, Till Sawala, Simon D. M. White, Richard Bower, Robert A. Crain, Michelle Furlong, Matthieu Schaller, Joop Schaye, and Tom Theuns. The unexpected diversity of dwarf galaxy rotation curves. *Monthly Notices of the Royal Astronomical Society*, 452(4):3650–3665, 2015. doi: 10.1093/mnras/stv1504.

[8] James S. Bullock and Michael Boylan-Kolchin. Small-scale challenges to the ΛCDM paradigm. *Annual Review of Astronomy and Astrophysics*, 55:343–387, 2017. doi: 10.1146/annurev-astro-091916-055313.

[9] Stacy S. McGaugh, Federico Lelli, and James M. Schombert. Radial acceleration relation in rotationally supported galaxies. *Physical Review Letters*, 117:201101, 2016. doi: 10.1103/PhysRevLett.117.201101.

[10] Benoit Famaey and Stacy S. McGaugh. Modified Newtonian dynamics (MOND): Observational phenomenology and relativistic extensions. *Living Reviews in Relativity*, 15(10), 2012. doi: 10.12942/lrr-2012-10.

[11] Tobias Mistele, Stacy S. McGaugh, Federico Lelli, James M. Schombert, and Pengfei Li. Indefinitely flat circular velocities and the baryonic Tully–Fisher relation from weak lensing. *Astrophysical Journal Letters*, 969:L3, 2024. doi: 10.3847/2041-8213/ad54b0.

[12] Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, 2005. doi: 10.1137/1.9780898717921.

[13] Andrew Gelman, Xiao-Li Meng, and Hal Stern. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, 6(4):733–807, 1996.

[14] George W. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992. doi: 10.1109/5.192069.

[15] J. Zico Kolter and Tommi Jaakkola. Approximate inference in additive factorial HMMs with application to energy disaggregation. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, volume 22 of *Proceedings of Machine Learning Research*, pages 1472–1482. PMLR, 2012.

[16] Daniel D. Lee and H. Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791, 1999. doi: 10.1038/44565.

[17] Michal Aharon, Michael Elad, and Alfred Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11):4311–4322, 2006. doi: 10.1109/TSP.2006.881199.

[18] Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro. Online learning for matrix factorization and sparse coding. *Journal of Machine Learning Research*, 11:19–60, 2010.

[19] Federico Lelli, Stacy S. McGaugh, and James M. Schombert. SPARC: Mass models for 175 disk galaxies with Spitzer photometry and accurate rotation curves. *Astronomical Journal*, 152:157, 2016. doi: 10.3847/0004-6256/152/6/157.

[20] Mordehai Milgrom. A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *Astrophysical Journal*, 270:365–370, 1983. doi: 10.1086/161130.

[21] Bruno A. Olshausen and David J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996. doi: 10.1038/381607a0.

[22] Pengfei Li, Federico Lelli, Stacy S. McGaugh, and James M. Schombert. A comprehensive catalog of dark matter halo models for SPARC galaxies. *Astrophysical Journal Supplement Series*, 247(1):31, 2020. doi: 10.3847/1538-4365/ab700e.

[23] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, 2020. doi: 10.1073/pnas.1912789117.

[24] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.