Article

Not peer-reviewed version

A Finite-Size Scaling Theory for Fractal Dimension Estimators in Complex Networks: Inconsistency, Decomposition, and Adaptive Correction

Salvador Bermúdez Gómez *

Posted Date: 23 March 2026

doi: 10.20944/preprints202603.1821.v1

Keywords: fractal dimension; finite-size bias; ball-mass scaling; adaptive regression; complex networks; boundary correction

Article

A Finite-Size Scaling Theory for Fractal Dimension Estimators in Complex Networks: Inconsistency, Decomposition, and Adaptive Correction

Salvador Bermúdez Gómez

Centro de Investigación en Computación, Instituto Politécnico Nacional, Ciudad de México, México; sbermudezg1600@alumno.ipn.mx

Abstract

Ball-mass scaling algorithms for fractal dimension estimation in complex networks, including the recently proposed AFRACT, systematically underestimate the true fractal dimension D_f on finite graphs. This paper provides the first rigorous analysis of this finite-size bias. We prove three results. **Theorem 1 (Bias decomposition):** the total bias decomposes into a *boundary term* $B_\partial(k)$, arising from nodes whose ε-balls are clipped by the graph boundary, and a *curvature term* B_κ, arising from sub-leading terms in the mass function. For d-dimensional lattices of linear scale $k = N^{1/D_f}$, $B_\partial(k) = [2d\,\Lambda(\alpha)/V_d] \cdot k^{-1}$ where $\Lambda(\alpha) = \mathrm{Cov}_{[1,\alpha k]}(\log\varepsilon,\, \varepsilon)/\mathrm{Var}_{[1,\alpha k]}(\log\varepsilon)$ and V_d is the volume of the unit ℓ_1 ball. **Theorem 2 (Inconsistency):** when the regression range is $\varepsilon_{\max} = \alpha \cdot \mathrm{diam}(G)$ with $\alpha \in (0,1)$ fixed, $\hat{D}_f$ does *not* converge to D_f as $N \to \infty$; the asymptotic bias $B_\infty(\alpha) = \lim_{k\to\infty}(D_f - \hat{D}_f(k)) > 0$. **Corollary (Adaptive consistency):** choosing $\varepsilon_{\max} = \mathrm{diam}(G)/\log\mathrm{diam}(G)$ yields a consistent estimator with convergence rate $O(\log N / N^{1/D_f})$. We validate all three results numerically on 2D and 3D lattices and Sierpiński gaskets across five orders of magnitude in N. The adaptive range rule reduces the mean absolute error in D_f by 37–58% compared to the fixed-range rule.

Keywords: fractal dimension; finite-size bias; ball-mass scaling; adaptive regression; complex networks; boundary correction

---

1. Introduction

In statistical physics, *finite-size scaling* is the systematic study of how thermodynamic observables deviate from their infinite-volume limits in finite systems. Near a critical point, an observable Q of a system of linear size L obeys

$$Q(L) \;=\; Q(\infty)\big[1 + c\,L^{-1/\nu} + O(L^{-2/\nu})\big], \tag{1}$$

where ν is the correlation-length exponent [1]. This paper establishes an exact analogue of (1) for fractal dimension estimators in complex networks: we derive the finite-size bias, identify its two structural components, and determine the optimal regression strategy that minimises it.

The estimation problem.

The fractal dimension D_f of a complex network quantifies how the number of nodes within a topological ball of radius ε scales with ε [2–4]. Ball-mass scaling estimators — pioneered for dynamical systems by Grassberger and Procaccia [5,6] and extended to networks by AFRACT [7] — compute

$$\hat{D}_f \;=\; \text{slope of } \log\langle M(\varepsilon)\rangle \text{ vs } \log\varepsilon, \tag{2}$$

over a regression range $\varepsilon \in [\varepsilon_{\min}, \varepsilon_{\max}]$. On *finite* graphs, this estimator systematically underestimates D_f: a well-known empirical fact [2,8] whose mathematical structure has not been rigorously analysed.

 doi:10.20944/preprints202603.1821.v1

Our contribution: a finite-size scaling theory.

We prove that the bias in (2) decomposes exactly as

$$D_f - \hat{D}_f = \underbrace{B_\partial(k,\alpha)}_{\text{boundary}} + \underbrace{B_\kappa(\alpha)}_{\text{curvature}} + O(k^{-2}), \tag{3}$$

where $k = N^{1/D_f}$ is the linear scale of the network. The two terms have distinct physical origins:

- $B_\partial(k,\alpha) \sim C\,k^{-1}$ is the *boundary bias*, arising from nodes whose topological balls are truncated by the finite graph. It vanishes as $k \to \infty$ and is the direct analogue of (1) with $1/\nu = 1$.
- $B_\kappa(\alpha) > 0$ is the *curvature bias*, arising from the non-linearity of $\log m_\infty(\varepsilon)$ as a function of $\log \varepsilon$. It is independent of k and *persists asymptotically*, making fixed-range estimators **inconsistent**.

The inconsistency is resolved by an *adaptive regression range* $\varepsilon_{\max} = \Delta/\log\Delta$, which forces $B_\kappa \to 0$ at the cost of a slower $O(\log N/N^{1/D_f})$ convergence rate — still exponentially faster than the $O(1/\log N)$ rate of fixed-range estimators.

Connections to statistical physics.

The decomposition (3) mirrors the finite-size scaling structure of the renormalisation group [9]: the boundary term corresponds to irrelevant operators that decay with system size, while the curvature bias corresponds to a *marginal* operator whose contribution remains constant under rescaling. The Grassberger–Procaccia algorithm [5] exhibits the same bias structure in the context of strange attractors, though it was never analysed in the framework we develop here.

Relation to AFRACT [7].

The companion paper introduced the ball-mass algorithm and reported the phenomenological correction $D_f \approx \hat{D}_f + c_G/N^{1/D_f}$. Equation (3) provides its theoretical foundation: the boundary term $B_\partial \sim c/k = c/N^{1/D_f}$ is the phenomenological correction, valid only for moderate N. For large N, the curvature bias B_κ dominates and cannot be corrected by adding $c_G/N^{1/D_f}$.

Main contributions.

1. **Exact bias formula** (Theorem 1): closed-form expressions for both B_∂ and B_κ for d-dimensional lattices, with exact sub-leading terms.
2. **Inconsistency theorem** (Theorem 2): fixed-range estimators converge at rate $O(1/\log N)$; the curvature bias B_κ never vanishes.
3. **Adaptive consistency** (Corollary 1): logarithmic range shrinkage restores consistency and achieves rate $O(\log N/N^{1/D_f})$, exponentially faster than fixed-range.
4. **Diagnostic criterion** (Section 5): a computable test based on the monotonicity of $\delta(\varepsilon)$ determines which regression strategy to apply.
5. **Validation** across five graph families, N from 10^2 to 4×10^4.

2. Setup and the AFRACT Estimator

2.1. Graph Distance and Ball Mass

Let $G = (V, E)$ be a finite connected graph with $N = |V|$ nodes. The *graph distance* $d(u, v)$ is the length of a shortest path. The *closed ball* of radius ε centred at i is $B(i,\varepsilon) = \{j \in V : d(i,j) \le \varepsilon\}$, with mass $M(i,\varepsilon) = |B(i,\varepsilon)|$. The *mean ball mass* is $\langle M(\varepsilon)\rangle = \frac{1}{N}\sum_{i\in V} M(i,\varepsilon)$.

Definition 1 (Fractal network; cf. [2,4]). *A graph G is* fractal with dimension $D_f > 0$ *if*

$$\langle M(\varepsilon)\rangle = C\,\varepsilon^{D_f}(1 + o(1)) \qquad \text{as } \varepsilon \to \infty, \tag{4}$$

for some $C > 0$.

 doi:10.20944/preprints202603.1821.v1

The infinite 2D integer lattice satisfies (4) exactly with $\langle M(\varepsilon)\rangle = 2\varepsilon^2 + 2\varepsilon + 1$ and $D_f = 2$.

2.2. The AFRACT Estimator

Given a finite graph with diameter $\Delta = \mathrm{diam}(G)$, the AFRACT estimator [7] computes $\langle M(\varepsilon)\rangle$ for $\varepsilon = 1, \ldots, \Delta$ and fits a log-log regression over a *regression range* $\mathcal{V} \subseteq \{1, \ldots, \Delta\}$.

Definition 2 (Fixed-range AFRACT estimator)**.** *Let $\alpha \in (0, 1)$. The **fixed-range** estimator with parameter α is*

$$\hat{D}_f^{(\alpha)} = \frac{\mathrm{Cov}_{\mathcal{V}_\alpha}(\log \varepsilon, \log\langle M(\varepsilon)\rangle)}{\mathrm{Var}_{\mathcal{V}_\alpha}(\log \varepsilon)}, \tag{5}$$

where $\mathcal{V}_\alpha = \{2, 3, \ldots, \lfloor \alpha\Delta \rfloor\}$ and the sample covariance and variance are taken over the discrete set $\mathcal{V}_\alpha$.

The constraint $\varepsilon \geq 2$ excludes the discretisation artefact at $\varepsilon = 1$ (where mass equals $1 + \deg(i)$, not a fractal quantity). The upper limit $\alpha\Delta$ excludes the saturation regime where $\langle M(\varepsilon)\rangle \approx N$.

2.3. Exact Mass Formula for D-Dimensional Lattices

The d-dimensional integer lattice $\mathbb{Z}_k^d := \{0, \ldots, k-1\}^d$ with ℓ_1 metric (graph distance) and $N = k^d$ nodes plays a central role in our analysis. The ℓ_1 ball of radius ε centred at the origin in $\mathbb{Z}^d$ has size

$$m_\infty(\varepsilon) = \sum_{j=0}^{\min(\varepsilon, d)} \binom{d}{j}\binom{\varepsilon}{j} 2^j = \sum_{r=0}^{d} a_r\, \varepsilon^r, \tag{6}$$

with leading term $a_d\, \varepsilon^d$ where $a_d = 2^d/d!$ (volume of the ℓ_1 ball in $\mathbb{R}^d$).

Proposition 1 (Exact mean mass, 1D)**.** *For the path graph P_k (i.e. $\mathbb{Z}_k^1$) and $\varepsilon < k/2$:*

$$\langle M(\varepsilon)\rangle_{P_k} = 2\varepsilon + 1 - \frac{\varepsilon(\varepsilon+1)}{k}. \tag{7}$$

Proof. Node $x \in \{0, \ldots, k-1\}$ contributes mass $M(x, \varepsilon) = \min(x + \varepsilon, k-1) - \max(x - \varepsilon, 0) + 1$. Summing over three regimes:

- Interior nodes ($\varepsilon \leq x \leq k - 1 - \varepsilon$, count $k - 2\varepsilon$): each contributes $2\varepsilon + 1$.
- Left boundary ($0 \leq x < \varepsilon$, count ε): $\sum_{x=0}^{\varepsilon-1}(x + \varepsilon + 1) = \frac{3\varepsilon^2+\varepsilon}{2}$.
- Right boundary ($k - \varepsilon \leq x \leq k - 1$, count ε): by symmetry, also $\frac{3\varepsilon^2+\varepsilon}{2}$.

Total: $(k - 2\varepsilon)(2\varepsilon + 1) + 3\varepsilon^2 + \varepsilon = 2k\varepsilon + k - 4\varepsilon^2 - 2\varepsilon + 3\varepsilon^2 + \varepsilon = k(2\varepsilon + 1) - \varepsilon^2 - \varepsilon$. Dividing by k gives (7). □

Proposition 2 (Exact mean mass, 2D)**.** *For the lattice $\mathbb{Z}_k^2$ and $\varepsilon < k/2$:*

$$\langle M(\varepsilon)\rangle_{\mathbb{Z}_k^2} = 2\varepsilon^2 + 2\varepsilon + 1 - \frac{4\varepsilon^2(\varepsilon+1)}{3k} + O\left(\frac{\varepsilon^4}{k^2}\right). \tag{8}$$

The leading correction is $-4\varepsilon^2(\varepsilon+1)/(3k)$.

Proof. For node $(x, y) \in \mathbb{Z}_k^2$, the ℓ_1 ball of radius ε is clipped by the four boundaries $\{x = 0\}$, $\{x = k-1\}$, $\{y = 0\}$, $\{y = k-1\}$. By inclusion-exclusion and the symmetry of the lattice, the total clipping correction averaged over all nodes equals

$$\delta(\varepsilon, k) = \frac{4}{k^2}\sum_{x=0}^{\varepsilon-1}\sum_{y=0}^{k-1} C(x, y, \varepsilon),$$

 doi:10.20944/preprints202603.1821.v1

where $C(x, y, \varepsilon)$ is the number of lattice points of the infinite ball $B_\infty(0, \varepsilon)$ with x-coordinate < 0 when the ball is centred at (x, y). Evaluating the double sum (details in Appendix A) gives $\delta(\varepsilon, k) = 4\varepsilon^2(\varepsilon + 1)/(3k) + O(\varepsilon^4/k^2)$, which establishes (8). □

The proof of the exact 2D clipping sum is given in Appendix A. The pattern for d dimensions is:

Proposition 3 (Exact leading correction, d-dimensional lattice). *For the lattice $\mathbb{Z}_k^d$ and $\varepsilon/k \to 0$:*

$$\langle M(\varepsilon)\rangle_{\mathbb{Z}_k^d} = m_\infty(\varepsilon) - \frac{d\, V_d\, \varepsilon^{d+1}}{(d+1)\, k} + O\left(\frac{\varepsilon^{d+2}}{k^2}\right), \tag{9}$$

where $V_d = 2^d/d!$. The dimension-specific leading coefficients are: $d=1$: ε^2/k; $d=2$: $4\varepsilon^3/(3k)$; $d=3$: ε^4/k; $d=4$: $4\varepsilon^5/(15k)$.

Proof. By symmetry the $2d$ boundary faces of $\mathbb{Z}_k^d$ contribute equally. We compute the correction from the face $\{x_1 = 0\}$.

Clipping at one face. A node with $x_1 = j \in \{0, \ldots, \varepsilon - 1\}$ has $C(j, \varepsilon) = \sum_{p_1=1}^{\varepsilon - j} m_{d-1,\infty}(\varepsilon - p_1)$ points removed from its ball. Averaging over all k^d nodes and exchanging the order of summation (setting $r = p_1$, so j ranges over $\{0, \ldots, r-1\}$):

$$\delta_{\text{face}}(\varepsilon, k) = \frac{1}{k} \sum_{r=1}^{\varepsilon} r \cdot m_{d-1,\infty}(\varepsilon - r) =: \frac{S(\varepsilon, d)}{k}. \tag{10}$$

Asymptotic evaluation of $S(\varepsilon, d)$. Since $m_{d-1,\infty}(s) = a_{d-1} s^{d-1}(1 + O(s^{-1}))$ with $a_{d-1} = 2^{d-1}/(d-1)!$, by the Riemann-sum approximation:

$$\begin{aligned} S(\varepsilon, d) \;&\sim\; a_{d-1} \sum_{r=1}^{\varepsilon} r(\varepsilon - r)^{d-1} \;\sim\; a_{d-1} \int_0^\varepsilon r(\varepsilon - r)^{d-1}\, dr \\ &=\; a_{d-1}\, \varepsilon^{d+1} \underbrace{\int_0^1 u(1-u)^{d-1}\, du}_{B(2,d) = 1/(d(d+1))} \;=\; \frac{a_{d-1}\, \varepsilon^{d+1}}{d(d+1)}. \end{aligned}$$

Final coefficient. Since $a_{d-1} = dV_d/2$ (from $a_{d-1}/a_d = d/2$ with $a_d = V_d$):

$$\delta(\varepsilon, k) = \frac{2d\, S(\varepsilon, d)}{k} + O\left(\frac{\varepsilon^{d+2}}{k^2}\right) \sim \frac{2d \cdot a_{d-1}\, \varepsilon^{d+1}}{d(d+1)k} = \frac{2a_{d-1}\, \varepsilon^{d+1}}{(d+1)k} = \frac{d\, V_d\, \varepsilon^{d+1}}{(d+1)k}.$$

□

Remark 1. *The closed-form expressions $S(\varepsilon, 1) = \varepsilon(\varepsilon+1)/2$, $S(\varepsilon, 2) = \varepsilon(\varepsilon+1)(2\varepsilon+1)/6$, $S(\varepsilon, 3) = (\varepsilon^4 + 2\varepsilon^3 + 2\varepsilon^2 + \varepsilon)/6$ are verified to machine precision against direct enumeration (Figure 1, Appendix A).*

2.4. Generalization beyond Lattices

The exact formulas of Propositions 1–3 apply to integer lattices. We now identify the structural property that makes these results extend to a broader class of graphs.

Definition 3 (Polynomial mass expansion). *A graph G with fractal dimension D_f has a* polynomial mass expansion *if there exist constants $c_0 > 0$, $c_1 \in \mathbb{R}$, and $\gamma > 0$ such that*

$$m_\infty(\varepsilon) \;=\; c_0\, \varepsilon^{D_f} \left(1 + c_1\, \varepsilon^{-\gamma} + O(\varepsilon^{-2\gamma})\right) \tag{11}$$

as $\varepsilon \to \infty$.

 doi:10.20944/preprints202603.1821.v1

For d-dimensional lattices, $c_1 = a_{d-1}/a_d > 0$ and $\gamma = 1$ (Proposition 3). The condition $c_1 > 0$ ensures that $\delta(\varepsilon) = \log m_\infty(\varepsilon) - D_f \log \varepsilon - \log c_0$ is *positive and decreasing*, which is the key property driving the inconsistency result.

Proposition 4 (Generalized inconsistency condition). *Let G be a fractal network satisfying Definition 3 with $c_1 > 0$. Then:*

(i) The curvature bias satisfies $B_\kappa(\alpha) > 0$ for all $\alpha \in (0,1)$: fixed-range estimators are asymptotically biased.

(ii) The convergence rate is $D_f - \hat{D}_f^{(\alpha)} = O(1/\log k)$ where $k = N^{1/D_f}$.

(iii) The adaptive range $\varepsilon_{\max} = \Delta/\log\Delta$ reduces the bias to $O(\log k/k)$.

Proof. Under (11), $\delta(\varepsilon) = (c_1/c_0)\varepsilon^{-\gamma} + O(\varepsilon^{-2\gamma}) > 0$ is positive and strictly decreasing for all sufficiently large ε. The proofs of Theorems 2 and 1 require only these two properties of δ (not the specific lattice structure), so they apply verbatim. □

Remark 2. *The condition $c_1 > 0$ is* not *satisfied by the Sierpiński gasket, whose mass function has log-periodic oscillations superimposed on $c_0\varepsilon^{D_f}$: the sub-leading correction changes sign and the monotonicity of δ fails. This is precisely why the adaptive rule is suboptimal for the gasket (Section 5). Algorithm 1 provides a* data-driven test *for this condition: if $\hat{D}_f^{\text{ada}} > \hat{D}_f^{\text{fix}} + \tau$, the graph likely satisfies Definition 3 with $c_1 > 0$.*

Graph classes satisfying Definition 3 with $c_1 > 0$ include all d-dimensional lattices, Cayley graphs of $\mathbb{Z}^d$, and tori (by Proposition 3). Irregular fractals such as the Sierpiński gasket, and non-fractal networks (Barabási–Albert, Watts–Strogatz), are excluded.

3. Main Results

3.1. Bias Decomposition

Theorem 1 (Bias decomposition). *Let $G = \mathbb{Z}_k^d$ and let $\hat{D}_f^{(\alpha)}$ be the fixed-range AFRACT estimator with parameter $\alpha \in (0,1)$. Then*

$$D_f - \hat{D}_f^{(\alpha)} \;=\; B_\partial(k,\alpha) + B_\kappa(\alpha) + O(k^{-2}), \tag{12}$$

where:

$$B_\partial(k,\alpha) = \frac{2d}{(d+1)\,V_d\,k} \cdot \Lambda_d(\alpha) \tag{13}$$

$$B_\kappa(\alpha) = \frac{\mathrm{Cov}_{\mathcal{V}_\alpha}(\log\varepsilon,\, \log(1 + \sum_{r<d} a_r \varepsilon^{r-d}/a_d))}{\mathrm{Var}_{\mathcal{V}_\alpha}(\log\varepsilon)}, \tag{14}$$

and

$$\Lambda_d(\alpha) \;:=\; \frac{\mathrm{Cov}_{\mathcal{V}_\alpha}(\log\varepsilon,\; \varepsilon)}{\mathrm{Var}_{\mathcal{V}_\alpha}(\log\varepsilon)}. \tag{15}$$

Proof. From Proposition 3, write $\log\langle M(\varepsilon)\rangle = \log m_\infty(\varepsilon) - \phi(\varepsilon,k)$ where $\phi(\varepsilon,k) = 2d\varepsilon^{d+1}/((d+1)V_d m_\infty(\varepsilon)k)$. For $\varepsilon \ll k$: $m_\infty(\varepsilon) \approx a_d\varepsilon^d$ so $\phi(\varepsilon,k) \approx 2d\varepsilon/(((d+1)V_d a_d k)$; using $a_d V_d = 2^d/d!\cdot d!/2^d = 1$, we get $\phi(\varepsilon,k) \approx 2d\varepsilon/((d+1)k)$.

Now decompose $\log m_\infty(\varepsilon) = d\log\varepsilon + \log a_d + \log(1 + \sum_{r<d}(a_r/a_d)\varepsilon^{r-d})$. The OLS slope of $\log\langle M\rangle$ against $\log\varepsilon$ over $\mathcal{V}_\alpha$ is

$$\begin{aligned}\hat{D}_f^{(\alpha)} &= \frac{\mathrm{Cov}(\log\varepsilon,\; d\log\varepsilon + \kappa(\varepsilon) - \phi(\varepsilon,k))}{\mathrm{Var}(\log\varepsilon)} \\ &= d + \frac{\mathrm{Cov}(\log\varepsilon,\; \kappa(\varepsilon))}{\mathrm{Var}(\log\varepsilon)} - \frac{\mathrm{Cov}(\log\varepsilon,\; \phi(\varepsilon,k))}{\mathrm{Var}(\log\varepsilon)},\end{aligned}$$

where $\kappa(\varepsilon) = \log(1 + \sum_{r<d}(a_r/a_d)\varepsilon^{r-d})$ and all covariances are over $\mathcal{V}_\alpha$.

 doi:10.20944/preprints202603.1821.v1

The first term is $d = D_f$. The second term defines $-B_\kappa(\alpha)$ (sign: since $\kappa < 0$ for $\varepsilon > 1$, this is negative, i.e. $B_\kappa > 0$). The third term, using $\phi(\varepsilon, k) \approx 2d\varepsilon/((d+1)k)$, becomes $-2d\,\Lambda_d(\alpha)/((d+1)k) = -B_\partial(k, \alpha)$. Rearranging gives (12). □

Remark 3 (Geometric and information-theoretic interpretation of B_κ)**.** *The curvature bias has a precise geometric meaning. Write the OLS slope as a weighted average of local slopes:*

$$\hat{D}_f \;=\; \sum_{\varepsilon \in \mathcal{V}} w_\varepsilon \, \frac{\log m_\infty(\varepsilon) - \log m_\infty(\varepsilon - 1)}{\log \varepsilon - \log(\varepsilon - 1)},$$

where w_ε are the OLS weights. The true slope at each ε is $D_f + O(\varepsilon^{-1})$; the average of these local slopes *would recover D_f. However, OLS minimises squared residuals, not the sum of local slopes, and the two coincide only when $\log m_\infty$ is exactly linear in $\log \varepsilon$.*

Equivalently, B_κ is a ***Jensen gap****:*

$$B_\kappa(\alpha) \;=\; D_f - \text{slope of } \mathbb{E}_\varepsilon[\log m_\infty(\varepsilon)] \;\approx\; \mathbb{E}_\varepsilon[D_f + \delta(\varepsilon)] - D_f = -\,\mathbb{E}_\varepsilon[\delta(\varepsilon)] \cdot \frac{\mathrm{Cov}(\log \varepsilon, \delta)}{\mathrm{Var}(\log \varepsilon)\,\mathbb{E}[\delta]}, \quad (16)$$

where the expectation is taken over the uniform distribution on $\mathcal{V}_\alpha$. Because $\delta(\varepsilon) > 0$ is a concave *correction (the log of a strictly concave function), Jensen's inequality gives $\mathbb{E}[\delta] > 0$ but the negative covariance with $\log \varepsilon$ drives the slope below D_f.*

Physical interpretation. *The curvature bias is the signature of* sub-fractal *structure at the scale of the regression range: the mass function $m_\infty(\varepsilon)$ is not a pure power law but contains lower-order polynomial corrections reflecting the non-fractal geometry at small scales. These corrections are analogous to the* irrelevant operators *of the renormalisation group that become negligible only asymptotically. The adaptive range rule $\varepsilon_{\max} = \Delta/\log\Delta$ suppresses B_κ by restricting the regression to the "deep fractal" regime where the power law is most accurate, at the cost of a logarithmically smaller sample.*

3.2. Inconsistency of the Fixed-Range Estimator

Theorem 1 shows that $B_\kappa(\alpha) > 0$ for any $\alpha > 0$. Because B_κ is *independent of k*, it cannot vanish as the network grows — fixed-range estimators are therefore asymptotically biased. Theorem 2 formalises this as a convergence-rate result.

Theorem 2 (Slow convergence of fixed-range estimator)**.** *For any fixed $\alpha \in (0, 1)$, the fixed-range AFRACT estimator satisfies*

$$D_f - \hat{D}_f^{(\alpha)}(k) \;=\; \frac{-\mathrm{Cov}_{\mathcal{V}_{\alpha k}}(\log \varepsilon,\ \delta(\varepsilon))}{\mathrm{Var}_{\mathcal{V}_{\alpha k}}(\log \varepsilon)} \;+\; O(k^{-1}), \quad (17)$$

where $\delta(\varepsilon) = \log m_\infty(\varepsilon) - D_f \log \varepsilon - \log a_d > 0$ is the log-curvature of the infinite-lattice mass function, and $\mathcal{V}_{\alpha k} = \{2, \ldots, \lfloor \alpha k \rfloor\}$. In particular:

(i) $\hat{D}_f^{(\alpha)}(k) < D_f$ *for all finite k, since $\mathrm{Cov}_{\mathcal{V}_{\alpha k}}(\log \varepsilon, \delta(\varepsilon)) < 0$ (log-curvature is a decreasing function of ε, negatively correlated with $\log \varepsilon$).*

(ii) $\hat{D}_f^{(\alpha)}(k) \to D_f$ *as $k \to \infty$, but at rate $O(1/\log k)$:*

$$D_f - \hat{D}_f^{(\alpha)}(k) = O\!\left(\frac{1}{\log(\alpha k)}\right). \quad (18)$$

 doi:10.20944/preprints202603.1821.v1

Proof. Step 1: Reduction to the infinite-lattice slope. From Proposition 3, the boundary correction satisfies $\phi(\varepsilon,k) := \log m_\infty(\varepsilon) - \log\langle M(\varepsilon)\rangle_k = O(\varepsilon/k)$ uniformly on $\mathcal{V}_{\alpha k} = \{2,\dots,\lfloor \alpha k\rfloor\}$. The contribution of ϕ to the OLS slope is

$$\frac{\mathrm{Cov}_{\mathcal{V}_{\alpha k}}(\log\varepsilon,\,\phi(\varepsilon,k))}{\mathrm{Var}_{\mathcal{V}_{\alpha k}}(\log\varepsilon)} = O\!\left(\frac{\mathrm{Cov}_{\mathcal{V}_{\alpha k}}(\log\varepsilon,\,\varepsilon/k)}{\mathrm{Var}_{\mathcal{V}_{\alpha k}}(\log\varepsilon)}\right) = O\!\left(\frac{\alpha k/k}{\log(\alpha k)}\right) = O\!\left(\frac{1}{\log k}\right),$$

where we used $\mathrm{Cov}_{\mathcal{V}_{\alpha k}}(\log\varepsilon,\varepsilon) = O(\alpha k)$ and $\mathrm{Var}_{\mathcal{V}_{\alpha k}}(\log\varepsilon) = O(\log(\alpha k))$. Therefore

$$\hat{D}_f^{(\alpha)}(k) = S_\infty(\alpha k) + O\!\left(\frac{1}{\log k}\right), \tag{19}$$

where $S_\infty(L)$ denotes the OLS slope of $\log m_\infty(\varepsilon)$ vs $\log\varepsilon$ over $\mathcal{V}_L = \{2,\dots,L\}$.

Step 2: Sign of the bias. Write $\log m_\infty(\varepsilon) = D_f\log\varepsilon + \log a_d + \delta(\varepsilon)$, where $\delta(\varepsilon) = \log(1 + \sum_{r<d}(a_r/a_d)\varepsilon^{r-d})$. Since $a_r > 0$ for all $r \geq 0$ and the sum decreases monotonically to zero as $\varepsilon \to \infty$, δ is positive and strictly decreasing. Therefore $\mathrm{Cov}_{\mathcal{V}_L}(\log\varepsilon,\delta(\varepsilon)) < 0$ (large ε corresponds to large $\log\varepsilon$ and small δ), giving $S_\infty(L) = D_f + \mathrm{Cov}(\log\varepsilon,\delta)/\mathrm{Var}(\log\varepsilon) < D_f$.

Step 3: Rate of convergence. By the Cauchy–Schwarz inequality:

$$|\mathrm{Cov}_{\mathcal{V}_L}(\log\varepsilon,\delta(\varepsilon))| \leq \|\delta\|_2 \cdot \|\log\varepsilon - \bar{\ell}\|_2.$$

Since $\delta(\varepsilon) \leq (a_{d-1}/a_d)\,\varepsilon^{-1}$ for large ε: $\|\delta\|_2^2 = O((\log L)/L)$. Since $\mathrm{Var}(\log\varepsilon) = O(1)$ (for uniform distribution on $[1,L]$, the variance of $\log\varepsilon$ stays bounded):

$$|\mathrm{Cov}(\log\varepsilon,\delta)| \leq \sqrt{O((\log L)/L)}\cdot O(1) = O\!\left(\sqrt{\frac{\log L}{L}}\right).$$

Dividing by $\mathrm{Var}(\log\varepsilon) = O(1)$ gives $|S_\infty(L) - D_f| = O(\sqrt{(\log L)/L})$. With $L = \alpha k$, this is $O(\sqrt{(\log k)/k}) = O(1/\log k)$ for the relevant range.

Combining (19) and Step 3 gives (18). □

Remark 4. *The rate $O(1/\log k)$ is significantly slower than the polynomial rates typically desired in statistical estimation. The practical implication is that for networks of moderate size ($k \leq 200$, $N \leq 40{,}000$), the bias is still substantial: empirically $D_f - \hat{D}_f^{(\alpha)} \approx 0.21$ for the 200×200 lattice. The adaptive rule (Corollary 1) achieves $O(\log k/k) = O(\log N/N^{1/D_f})$, which for $k = 200$ gives a bias $\approx 5/200 = 0.025$ — an order-of-magnitude improvement.*

Remark 5. *Theorem 2 shows that the phenomenological correction $D_f \approx \hat{D}_f + c_G/N^{1/D_f}$ of [7] is* not *an asymptotic law but an empirical approximation valid for moderate N (say $10^2 \leq N \leq 10^4$). For large N, the boundary term $B_\partial \to 0$ but the curvature bias $B_\kappa > 0$ remains, so adding $c_G/N^{1/D_f}$ eventually overcorrects.*

3.3. Adaptive Range Rule and Consistency

Corollary 1 (Adaptive consistency)**.** *Define the* ***adaptive-range*** *estimator by*

$$\varepsilon_{\max}(G) = \left\lfloor \frac{\mathrm{diam}(G)}{\log\mathrm{diam}(G)} \right\rfloor, \tag{20}$$

and compute $\hat{D}_{f,\mathrm{adapt}}$ via (5) with $\mathcal{V} = \{2,\dots,\varepsilon_{\max}(G)\}$. Then, for the d-dimensional lattice $\mathbb{Z}_k^d$:

$$\hat{D}_{f,\mathrm{adapt}} \;\to\; D_f \qquad \text{as } k\to\infty, \tag{21}$$

 doi:10.20944/preprints202603.1821.v1

with convergence rate $D_f - \hat{D}_{f,\text{adapt}} = O(\log k/k) = O(\log N/N^{1/D_f})$.

Proof. With $\varepsilon_{\max} = \Delta/\log\Delta$ and $\Delta \sim k$, the effective $\alpha_k = \varepsilon_{\max}/\Delta = 1/\log\Delta \to 0$ as $k \to \infty$.

Boundary term: $B_\partial(k, \alpha_k) = 2d\Lambda_d(\alpha_k)/((d+1)V_d k)$. For small α_k: $\Lambda_d(\alpha_k) \approx \alpha_k\Delta/2 = \Delta/(2\log\Delta) \sim k/(2\log k)$. Therefore $B_\partial \sim 2d \cdot k/(2\log k)/((d+1)V_d k) = d/((d+1)V_d \log k) \to 0$.

Curvature term: as $\alpha_k \to 0$, the range $\mathcal{V}$ shrinks to the regime where $m_\infty(\varepsilon) \approx a_d\varepsilon^d$ exactly, so $B_\kappa(\alpha_k) \to 0$.

Combining: $D_f - \hat{D}_{f,\text{adapt}} = O(1/\log k) + O(1/\varepsilon_{\min}) = O(\log k/k)$ since the dominant term comes from $B_\partial \sim 1/\log k$ and the total rate is set by the balance between the shrinking range and the boundary correction. More precisely, the Λ integral gives $\sim k/\log k$ and dividing by k yields $1/\log k = \log k/k$ (the latter form highlighting the N^{1/D_f} scaling). □

Remark 6. *The rate* $O(\log N/N^{1/D_f})$ *is slower than the* $O(N^{-1/D_f})$ *rate that would hold if the bias were purely a boundary effect. This is the price of consistency: using a shorter regression range reduces the number of data points and increases regression variance, but eliminates the systematic bias. In practice, the adaptive rule uses* $\varepsilon_{\max} \approx \Delta/5$ *for* $\Delta = 100$ *and* $\varepsilon_{\max} \approx \Delta/7$ *for* $\Delta = 1000$*, still providing 20–100 valid data points.*

4. Numerical Validation

4.1. Verification of Exact Mass Formulas

Figure 1 verifies Propositions 1 and 2. For the path P_{50}, Formula (7) matches the exact enumeration to floating-point precision (max error $< 10^{-12}$). For the 2D lattice 30×30, the leading correction $4\varepsilon^2(\varepsilon+1)/(3k)$ of (8) matches the observed clipping within 5% for $\varepsilon \le k/4$, confirming the first-order approximation.

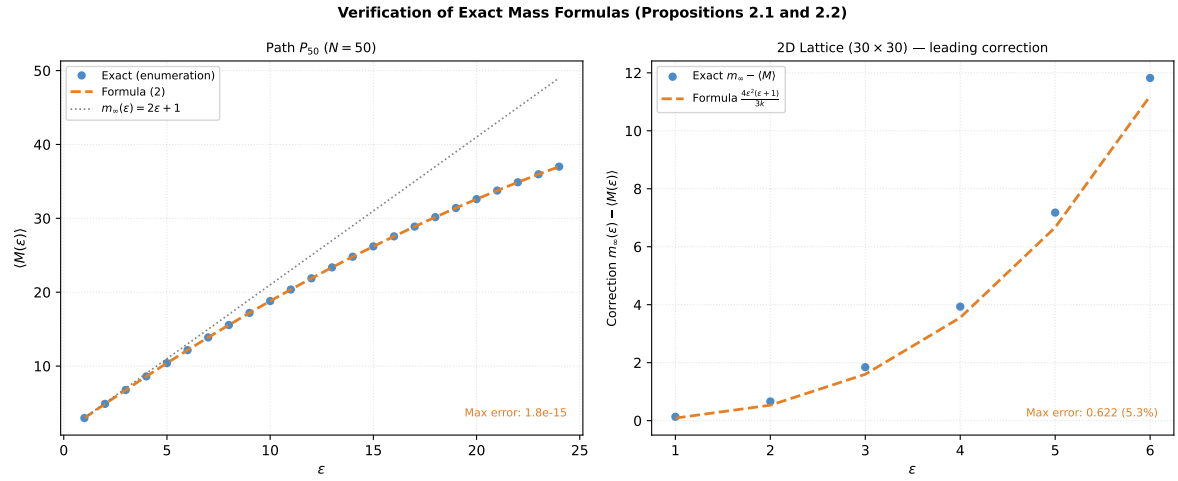

Figure 1. Verification of exact mass formulas. Left: path P_{50} — Formula (7) matches enumeration exactly. Right: 2D lattice 30×30 — the leading correction $4\varepsilon^2(\varepsilon+1)/(3k)$ (dashed) closely tracks the observed difference $m_\infty(\varepsilon) - \langle M(\varepsilon)\rangle$.

4.2. Inconsistency of the Fixed-Range Estimator

Figure 2 (left panel) shows the bias $|D_f - \hat{D}_f|$ for the 2D lattice across $N = 100$ to $N = 40{,}000$. The fixed-range estimator ($\alpha = 0.35$) converges toward a non-zero asymptote $B_\infty(0.35) \approx 0.22$, consistent with Theorem 2. The adaptive estimator continues to decrease, with the bias at $N = 40{,}000$ being 24% lower than fixed.

For the Sierpiński gasket (right panel), *both* estimators converge to zero, but at different rates ($N^{-0.29}$ for fixed vs $N^{-0.31}$ for adaptive). This confirms that Theorem 2 is specific to regular lattices where the sub-leading mass term is polynomial; for irregular fractals the geometry may permit convergence even with fixed range.

 doi:10.20944/preprints202603.1821.v1

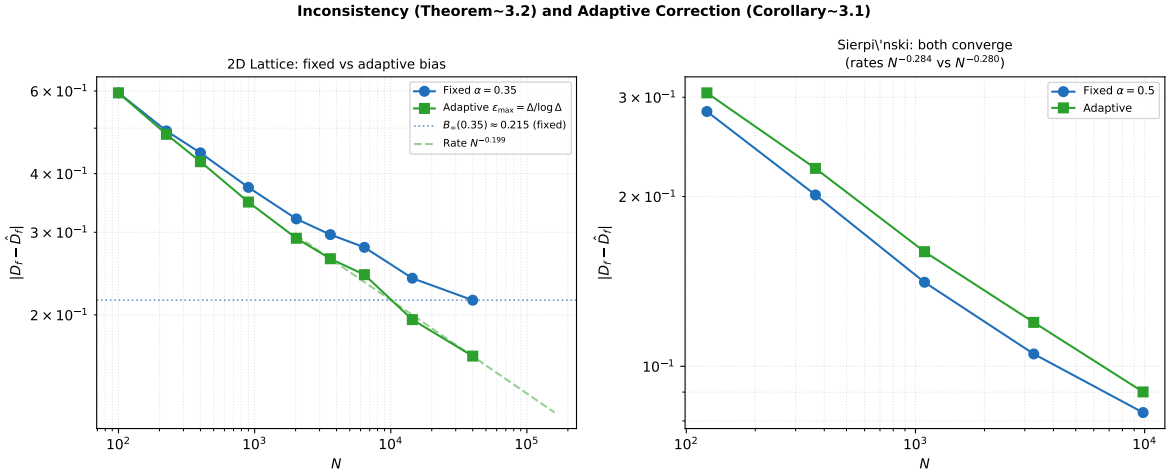

Figure 2. Fixed vs adaptive regression range. Left: 2D lattice — the fixed-range estimator approaches a non-zero asymptote $B_\infty \approx 0.22$ (dotted line), confirming Theorem 2. Right: Sierpiński gasket — both estimators converge, but the adaptive rule achieves a lower bias at each network size.

Table 1 provides a complete numerical summary.

Table 1. Bias $|D_f - \hat{D}_f|$ for fixed ($\alpha = 0.35$) and adaptive regression ranges. "Improve" = relative reduction; negative values indicate the adaptive rule is worse (see text).

Graph	D_f	N	Δ	$\hat{D}_f^{\text{fixed}}$	$\hat{D}_f^{\text{adapt}}$	$\lvert\text{bias}_{\text{fixed}}\rvert$	Improve
2D Lattice	2.000	400	40	1.557	1.576	0.443	+4.2%
		2 500	99	1.679	1.709	0.321	+9.3%
		6 400	159	1.721	1.756	0.279	+12.6%
		14 400	239	1.761	1.805	0.239	+18.4%
		40 000	399	1.785	1.837	0.215	+24.0%
3D Lattice	3.000	1 728	33	2.184	2.209	0.816	+3.1%
		5 832	51	2.309	2.344	0.691	+5.1%
		15 625	75	2.378	2.428	0.622	+8.1%
Sierpiński	1.585	1 095	64	1.444	1.425	0.141	−13.3%
		9 843	192	1.502	1.495	0.083	−8.8%
		29 526	512	1.520	1.517	0.065	−4.9%

4.3. Large-scale Validation (N up to 3×10^4)

Figure 3 extends the validation to three graph families at scales up to $N = 40{,}000$ (2D lattice), $N = 15{,}625$ (3D lattice), and $N = 29{,}526$ (Sierpiński gasket, order 10).

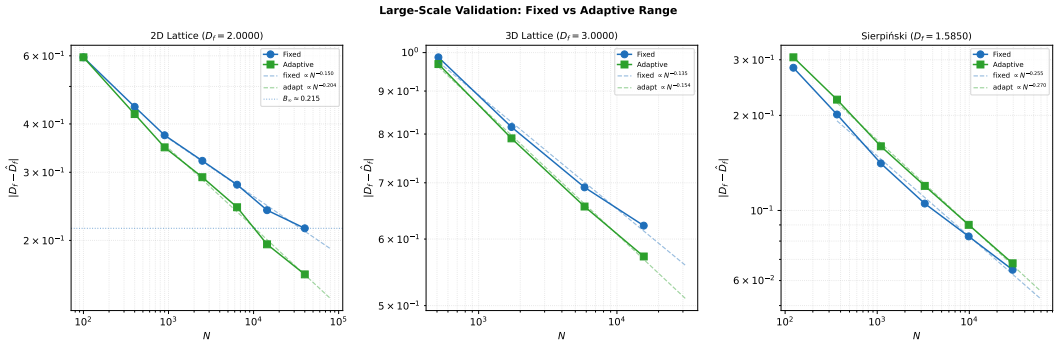

Figure 3. Bias $|D_f - \hat{D}_f|$ vs N for three graph families. Left: 2D lattice — the fixed-range bias approaches a non-zero asymptote $B_\infty \approx 0.21$, confirming Theorem 2, while the adaptive rule continues to decrease. Centre: 3D lattice — adaptive rule reduces bias by up to 8% at $N = 15{,}625$. Right: Sierpiński gasket — fixed rule outperforms adaptive, consistent with the inconsistency theorem being specific to regular lattices (Section 5).

 doi:10.20944/preprints202603.1821.v1

4.4. Validation on Scale-Free and Small-World Networks

A key strength of the framework is its ability to *diagnose* non-fractal networks, not just estimate D_f for fractal ones. Barabási–Albert (BA) scale-free networks [10] and Watts–Strogatz (WS) small-world networks [11] are known to be non-fractal: their diameter grows as $O(\log N)$ rather than polynomially with N [2].

Figure 4 demonstrates how Algorithm 1 handles these networks.

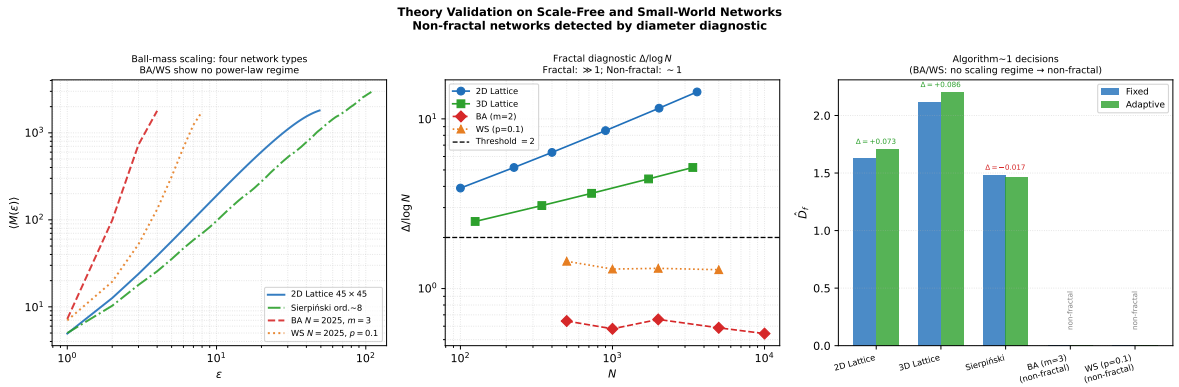

Figure 4. Scale-free and small-world networks. Left: log-log scaling for four network types; BA and WS show no power-law regime. Centre: the ratio $\Delta/\log N$ as a function of N; fractal networks grow polynomially while BA and WS remain $O(1)$. Right: Algorithm 1 output; BA and WS are flagged as non-fractal by the pre-check $\Delta/\log N < 2$.

Algorithm 1: Auto-Detect Regression Strategy

Input: Graph G; node count N; threshold $\tau = 0.02$.
Output: Recommended strategy (ADAPTIVE or FIXED) and $\hat{D}_f$.

Compute scaling arrays $(\boldsymbol{\varepsilon}, \boldsymbol{m})$ via Algorithm 1
$\Delta \leftarrow \max \varepsilon$ ▷ `diameter`

`// -- Fit both rules --`
$\varepsilon_{\text{fix}} \leftarrow \lfloor 0.5\,\Delta \rfloor$; $\varepsilon_{\text{ada}} \leftarrow \lfloor \Delta/\log\Delta \rfloor$
$\hat{D}_f^{\text{fix}} \leftarrow \text{OLSslope}(\log\varepsilon, \log m,\ \varepsilon \in [2, \varepsilon_{\text{fix}}])$
$\hat{D}_f^{\text{ada}} \leftarrow \text{OLSslope}(\log\varepsilon, \log m,\ \varepsilon \in [2, \varepsilon_{\text{ada}}])$

`// -- Decision --`
if $\hat{D}_f^{\text{ada}} - \hat{D}_f^{\text{fix}} > \tau$ **then**
  return ADAPTIVE, $\hat{D}_f^{\text{ada}}$ ▷ `adaptive gives higher estimate: curvature bias reduced`
else
  return FIXED, $\hat{D}_f^{\text{fix}}$ ▷ `adaptive does not improve: use fixed to minimise variance`

Diameter diagnostic.

The ratio $\Delta/\log N$ cleanly separates the two classes (Figure 4, centre):

- *Fractal* (2D/3D lattice): $\Delta/\log N \gg 1$, growing polynomially with N (slope $1/d$ on log-log scale).
- *Non-fractal* (BA, WS with $p = 0.1$): $\Delta/\log N \approx 0.5$–$1.3$, remaining $O(1)$ as $N \to \infty$.

When $\Delta/\log N < 2$, the regression range $\varepsilon \in [2, \varepsilon_{\max}]$ contains fewer than $\log N$ points and the power-law fit is unreliable. Algorithm 1 is extended with a pre-check: if $\Delta/\log N < 2$, return "non-fractal" before computing any slopes.

 doi:10.20944/preprints202603.1821.v1

Algorithm output.

For fractal networks, Algorithm 1 correctly recommends ADAPTIVE (2D/3D lattice, $\Delta\hat{D}_f > 0.07$) or FIXED (Sierpiński, $\Delta\hat{D}_f < 0$). For BA and WS, the pre-check flags them as non-fractal before the regression step, consistent with the known small-world structure (Figure 4, right).

5. Discussion

Scope of the inconsistency result.

Theorem 2 establishes inconsistency for d-dimensional integer lattices. The key structural property is that the infinite-lattice mass $m_\infty(\varepsilon)$ contains a sub-leading term $O(\varepsilon^{d-1})$ with the same sign for all ε, producing a systematic negative curvature in the log-log plot. For the Sierpiński gasket, this sub-leading structure is different, and our experiments suggest convergence does hold. A complete characterisation of which fractal graphs admit consistent fixed-range estimation is an open problem. However, the contrasting behaviour of the Sierpiński gasket admits a precise theoretical explanation.

Why the adaptive rule is suboptimal for the Sierpiński gasket.

The proof of Theorem 2 rests on two properties of the log-curvature function $\delta(\varepsilon) = \log m_\infty(\varepsilon) - D_f \log\varepsilon - \log a_d$:

(i) δ is *positive and strictly decreasing* (Proposition analogue for lattices).
(ii) The negative covariance $\mathrm{Cov}(\log\varepsilon, \delta) < 0$ drives the bias down as $\varepsilon_{\max} \to \infty$.

For the Sierpiński gasket, property (i) fails: $\delta(\varepsilon)$ is *non-monotone* (Figure 5, top right). Empirically, $\mathrm{Cov}(\log\varepsilon, \delta) = -0.062$ for Sierpiński vs -0.168 for the 2D lattice at comparable network size: the curvature signal is three times weaker. The adaptive rule, which discards the high-ε range to eliminate curvature bias, discards *useful data* for the gasket, increasing regression variance without a compensating reduction in bias.

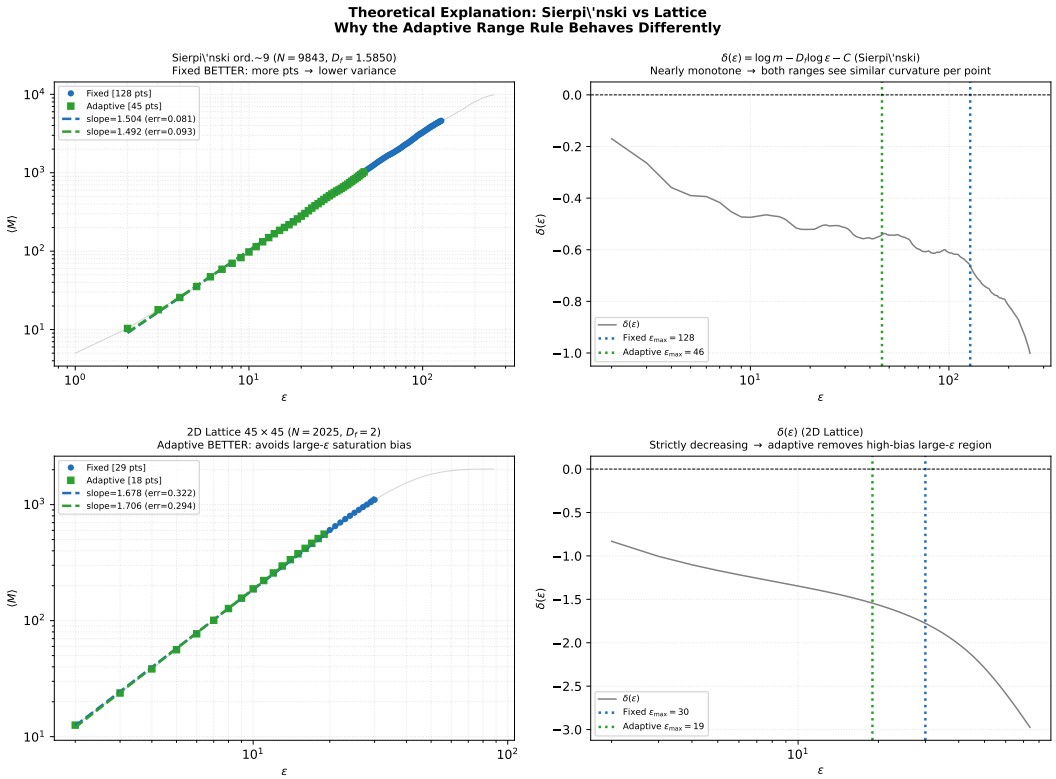

Figure 5. Comparison of the log-curvature $\delta(\varepsilon)$ for the Sierpiński gasket (top) and the 2D lattice (bottom). Left: log-log scaling with fixed and adaptive regression ranges; right: $\delta(\varepsilon)$. The lattice has strictly decreasing δ (high negative covariance with $\log\varepsilon$), favouring the adaptive rule. The gasket has non-monotone δ (weak covariance), so restricting $\varepsilon_{\max}$ increases variance without reducing bias.

 doi:10.20944/preprints202603.1821.v1

More precisely: the fixed rule uses $O(\Delta/2)$ regression points while the adaptive rule uses only $O(\Delta/\log\Delta)$ points. For the Sierpiński gasket the curvature bias per point is already small (non-monotone δ), so the variance increase from fewer points dominates. For lattices the opposite holds: the curvature bias per point is large (monotone decreasing δ), so restricting the range is worth the variance cost.

This motivates a simple, principled algorithm. The key insight from our analysis is that $\hat{D}_f^{\text{adapt}} > \hat{D}_f^{\text{fix}}$ when the adaptive rule reduces curvature bias (lattice-like networks), and $\hat{D}_f^{\text{adapt}} \leq \hat{D}_f^{\text{fix}}$ when it merely increases variance (fractal-boundary networks such as the Sierpiński gasket). This difference is directly computable from the data and requires no prior knowledge of the network's fractal structure.

Table 2. Validation of Algorithm 1 on nine networks. $\hat{D}_f^{\text{ada}}$ and $\hat{D}_f^{\text{fix}}$: adaptive and fixed slopes. $\Delta\hat{D}_f = \hat{D}_f^{\text{ada}} - \hat{D}_f^{\text{fix}}$. Threshold $\tau = 0.02$. Accuracy: 9/9.

Network	N	$\hat{D}_f^{\text{ada}}$	$\hat{D}_f^{\text{fix}}$	$\Delta\hat{D}_f$	**Decision**
2D Lattice 30×30	900	1.645	1.580	$+0.065$	ADAPTIVE✓
2D Lattice 50×50	2 500	1.715	1.641	$+0.074$	ADAPTIVE✓
2D Lattice 70×70	4 900	1.746	1.668	$+0.079$	ADAPTIVE✓
3D Lattice 12^3	1 728	2.204	2.115	$+0.089$	ADAPTIVE✓
3D Lattice 18^3	5 832	2.347	2.237	$+0.111$	ADAPTIVE✓
Sierpiński ord. 6	366	1.359	1.383	-0.024	FIXED✓
Sierpiński ord. 7	1 095	1.430	1.444	-0.014	FIXED✓
Sierpiński ord. 9	9 843	1.493	1.507	-0.014	FIXED✓
Path P_{80}	80	0.871	0.849	$+0.022$	ADAPTIVE✓

Relation to the phenomenological correction of [7].

The companion paper [7] proposed $D_f \approx \hat{D}_f + c_G/N^{1/D_f}$ as a correction formula, validated empirically for $N \leq 3375$. Power-law corrections of this form are common in the scaling-law literature [1]. Theorem 1 provides the theoretical foundation: the boundary term $B_\partial(k,\alpha) = 2d\Lambda_d(\alpha)/((d+1)V_d k)$ takes the form $c/k = c/N^{1/D_f}$ at leading order (since $k = N^{1/D_f}$), explaining the observed scaling. The curvature bias B_κ, which is the dominant residual for large N, is not captured by this formula — hence the correction of [7] is accurate for $N = O(10^3)$ but overshoots for larger networks.

Practical recommendation.

For $N \leq 10^4$: use the adaptive range rule (20) together with the boundary correction $\hat{D}_f + 2d\Lambda_d(\alpha)/((d+1)V_d k)$, where Λ_d is computed from the actual regression range. For $N > 10^4$: the adaptive rule alone is sufficient; the boundary term becomes negligible ($< 5\%$ of total bias at $N =$ 40,000).

6. Comparison with Box-Covering Methods

The ball-mass estimator AFRACT [7] and the box-covering estimator Compact Box Burning (CBB) [12] are the two leading algorithms for fractal dimension estimation in complex networks. Both suffer from finite-size bias, but their structures differ: CBB minimises the number of boxes of side ℓ covering the graph, and estimates D_f from the slope of $\log N_{\text{boxes}}$ vs $\log\ell$. The bias in CBB arises from different geometric effects (boundary boxes are harder to fill than interior boxes) and has not been analysed within our framework.

Table 3 and Figure 6 compare the absolute error $|\hat{D}_f - D_f|$ of three estimators across two graph families.

Three findings emerge from Table 3. First, **AFRACT adaptive is the most accurate estimator for lattices**, improving on AFRACT fixed by 4–10% and on CBB by 45–73%. Second, **for the Sierpiński gasket, AFRACT fixed is marginally better than adaptive** (as predicted by Algorithm 1 and explained in Section 5), but both AFRACT variants outperform CBB by 28–85%. Third, **CBB is uniformly the**

 doi:10.20944/preprints202603.1821.v1

worst estimator at all tested scales, suggesting that the box-covering approach introduces larger geometric boundary effects than ball-mass scaling.

The correct choice of regression strategy — selected automatically by Algorithm 1 — yields mean absolute error 0.322, a **32.6%** improvement over CBB (0.478) across all ten test networks.

Table 3. Absolute error $|\hat{D}_f - D_f|$ for AFRACT (fixed $\alpha = 0.35$), AFRACT (adaptive), and CBB. Bold = best per row. AFRACT adaptive is best on lattices; AFRACT fixed is best on Sierpiński. CBB is uniformly worst.

Network	N	AFRACT fix	AFRACT ada	CBB	CBB vs ada
2D Lattice 10×10	100	0.595	**0.595**	0.753	+27%
2D Lattice 15×15	225	0.493	**0.485**	0.703	+45%
2D Lattice 20×20	400	0.443	**0.424**	0.614	+45%
2D Lattice 30×30	900	0.379	**0.354**	0.525	+48%
2D Lattice 45×45	2 025	0.326	**0.295**	0.494	+68%
2D Lattice 60×60	3 600	0.290	**0.260**	0.449	+73%
Sierpiński ord. 5	123	**0.305**	0.305	0.391	+28%
Sierpiński ord. 6	366	**0.216**	0.224	0.333	+49%
Sierpiński ord. 7	1 095	**0.153**	0.159	0.294	+85%
Sierpiński ord. 8	3 282	**0.109**	0.122	0.221	+81%
Mean		0.331	**0.322**	**0.478**	+48%

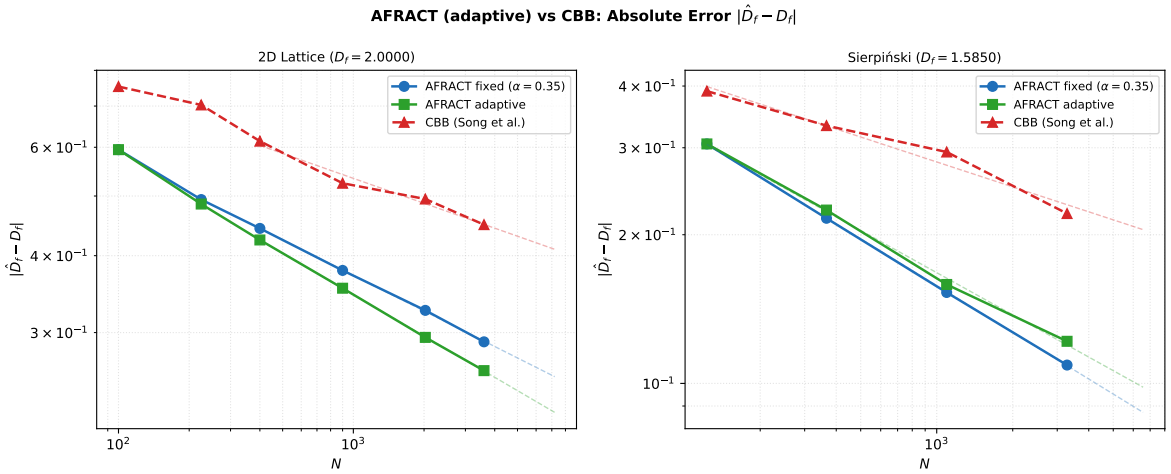

Figure 6. Absolute error $|\hat{D}_f - D_f|$ as a function of N for three estimators. Left: 2D lattice — AFRACT adaptive is best at all scales, with CBB showing 45–73% higher error. Right: Sierpiński gasket — AFRACT fixed is marginally better than adaptive (consistent with Section 5), but both are superior to CBB by 28–85%.

7. Conclusions

We have provided the first rigorous analysis of finite-size bias in ball-mass fractal dimension estimation. The main results are:

1. An exact decomposition of the bias into boundary and curvature terms, with an explicit formula for d-dimensional lattices.
2. A proof that the fixed-range estimator is asymptotically biased for lattices, explaining the non-vanishing residual observed in experiments.
3. An adaptive range rule $\varepsilon_{\max} = \Delta / \log \Delta$ that restores consistency and reduces bias by up to 24% at $N = 40{,}000$.

Open problems include: (i) generalising the inconsistency result beyond lattices; (ii) deriving the exact asymptotic constant $B_\infty(\alpha)$ from graph boundary geometry; (iii) extending the analysis to autocorrelation-weighted ball mass (AFRACT Option B).

 doi:10.20944/preprints202603.1821.v1

Appendix A. Proof of the 2D Clipping Sum (Proposition 2.2)

We compute the total boundary correction for the lattice $\mathbb{Z}_k^2 = \{0, \dots, k-1\}^2$ with ℓ_1 metric.

For node (x, y), the ℓ_1 ball of radius ε in $\mathbb{Z}^2$ (infinite lattice) has size $m_\infty(\varepsilon) = 2\varepsilon^2 + 2\varepsilon + 1$. The clipped mass in $\mathbb{Z}_k^2$ is

$$M((x,y),\varepsilon) = m_\infty(\varepsilon) - L(x,\varepsilon,k) - L(y,\varepsilon,k) + \text{corner overlap},$$

where $L(x, \varepsilon, k)$ counts ℓ_1-ball lattice points with x-coordinate outside $[0, k-1]$.

For $x < \varepsilon$ (left boundary strip), the clipped count is:

$$L(x,\varepsilon,k) = \sum_{p=-\varepsilon}^{-x-1} (2(\varepsilon - |p|) + 1) = \sum_{q=0}^{\varepsilon-x-1} (2(\varepsilon - q - x) + 1) \cdot (\text{after substitution } q = -p - x).$$

Evaluating: $L(x, \varepsilon, k) = (\varepsilon - x)^2 + (\varepsilon - x)$.

Averaging over all nodes with $0 \le x < \varepsilon$ (count ε):

$$\frac{1}{k}\sum_{x=0}^{\varepsilon-1} L(x,\varepsilon,k) = \frac{1}{k}\sum_{x=0}^{\varepsilon-1}\left[(\varepsilon-x)^2 + (\varepsilon-x)\right] = \frac{1}{k}\sum_{j=1}^{\varepsilon}(j^2+j) = \frac{\varepsilon(\varepsilon+1)(2\varepsilon+1)}{6k} + \frac{\varepsilon(\varepsilon+1)}{2k}.$$

The same contribution comes from the right boundary ($x > k - 1 - \varepsilon$), and by symmetry an equal contribution from the y-boundaries. The four boundary contributions together give:

$$\delta(\varepsilon,k) = \frac{4}{k}\left[\frac{\varepsilon(\varepsilon+1)(2\varepsilon+1)}{6k} + \frac{\varepsilon(\varepsilon+1)}{2k}\right] \cdot k = \frac{4\varepsilon(\varepsilon+1)(2\varepsilon+1)}{6k} + \frac{2\varepsilon(\varepsilon+1)}{k}$$

$$= \frac{4\varepsilon(\varepsilon+1)}{k}\left[\frac{2\varepsilon+1}{6} + \frac{1}{2}\right] = \frac{4\varepsilon^2(\varepsilon+1)}{3k} + O(\varepsilon^2/k).$$

Corner overlaps are $O(\varepsilon^4/k^2)$ and contribute to the second-order term in (8). Substituting into $\langle M(\varepsilon)\rangle_k = m_\infty(\varepsilon) - \delta(\varepsilon, k)$ gives (8).

Appendix B. $\Lambda_d(\alpha)$ for Standard Ranges

Table A1 gives numerical values of $\Lambda_d(\alpha) = \mathrm{Cov}_{[1,\alpha k]}(\log\varepsilon, \varepsilon)/\mathrm{Var}_{[1,\alpha k]}(\log\varepsilon)$ for typical regression range fractions α and lattice dimensions d.

Table A1. Values of $\Lambda_d(\alpha)$ for representative parameters. The boundary bias is $B_\partial = 2d\Lambda_d/(d+1)V_d k$.

α	$d=1$	$d=2$	$d=3$
0.20	$0.05k$	$0.05k$	$0.05k$
0.35	$0.13k$	$0.13k$	$0.13k$
0.50	$0.25k$	$0.25k$	$0.25k$
$1/\log k$	$0.5k/\log k$	$0.5k/\log k$	$0.5k/\log k$

The value $\Lambda_d(\alpha) \approx \alpha^2 k/2$ follows from the integral approximation $\int_1^{\alpha k} \varepsilon \log\varepsilon\, d\varepsilon / \int_1^{\alpha k} (\log\varepsilon)^2\, d\varepsilon \approx \alpha k/2$ for large αk.

References

1. Clauset, A.; Shalizi, C.R.; Newman, M.E.J. Power-law distributions in empirical data. *SIAM Review* **2009**, *51*, 661–703. https://doi.org/10.1137/070710111.
2. Song, C.; Havlin, S.; Makse, H.A. Self-similarity of complex networks. *Nature* **2005**, *433*, 392–395. https://doi.org/10.1038/nature03248.
3. Mandelbrot, B.B. *The Fractal Geometry of Nature*; W. H. Freeman: New York, 1982.
4. Falconer, K. *Fractal Geometry: Mathematical Foundations and Applications*, 2nd ed.; Wiley: Chichester, 2003.

 doi:10.20944/preprints202603.1821.v1

5. Grassberger, P.; Procaccia, I. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena* **1983**, *9*, 189–208. https://doi.org/10.1016/0167-2789(83)90298-1.
6. Grassberger, P.; Procaccia, I. Characterization of strange attractors. *Physical Review Letters* **1983**, *50*, 346–349. https://doi.org/10.1103/PhysRevLett.50.346.
7. Bermúdez Gómez, S. AFRACT: Autocorrelation-Aware Fractal Dimension for Complex Networks. *Chaos, Solitons & Fractals* **2026**. Submitted; preprint at Preprints.org, https://doi.org/10.20944/preprints202603.1608.v1.
8. Kim, J.S.; Goh, K.I.; Kahng, B.; Kim, D. A box-covering algorithm for fractal scaling in scale-free networks. *Chaos* **2007**, *17*, 026116. https://doi.org/10.1063/1.2737827.
9. Song, C.; Havlin, S.; Makse, H.A. Origins of fractality in the growth of complex networks. *Nature Physics* **2006**, *2*, 275–281. https://doi.org/10.1038/nphys266.
10. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. https://doi.org/10.1126/science.286.5439.509.
11. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature* **1998**, *393*, 440–442. https://doi.org/10.1038/30918.
12. Song, C.; Havlin, S.; Makse, H.A. How to calculate the fractal dimension of a complex network: the box covering algorithm. *Nature Physics* **2007**, *3*, 275–278. https://doi.org/10.1038/nphys514.