Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Mia C. Morrell March 9, 2020

Spatial and temporal criticality emerge from latent-variable models

by

Mia C. Morrell

Ilya M. Nemenman
Adviser

Audrey J. Sederberg
Co-Adviser

Department of Physics

Ilya M. Nemenman

Adviser

Audrey J. Sederberg

Co-Adviser

Gordon Berman

Committee Member

Stefan Boettcher

Committee Member

2020

Spatial and temporal criticality emerge from latent-variable models

By

Mia C. Morrell

Ilya M. Nemenman

Adviser

Audrey J. Sederberg

Co-Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Physics

2020

Abstract

Spatial and temporal criticality emerge from latent-variable models
By Mia C. Morrell

Understanding activity of large populations of neurons is difficult due to the combinatorial complexity of possible cell-cell interactions. A remedy is to note that many systems may be macroscopically described with models simpler than the system's microscopic behavior. This has been probed via a coarse-graining procedure on experimental neural recordings, which shows over two decades of scaling in free energy, variance, eigenvalue spectra, and correlation time [1], hinting that a mouse hippocampus operates in a critical regime. We investigated whether this scaling behavior could be explained as a result of coupling the neural population to latent dynamic stimuli. We conducted simulations of conditionally independent binary neurons coupled to a small number of long-timescale stochastic fields with and without periodic spatial stimuli (depicting neural place cells) and replicated the coarse-graining shown in [1]. In a biologically relevant regime, we find that much of the observed scaling [1] may be recreated by this model. This suggests that aspects of the scaling may be explained by coupling to hidden dynamic processes, a ubiquitous trait of neural systems.

[1] L. Meshulam et al. arXiv:1812.11904 [physics.bio-ph], 2019

Spatial and temporal criticality emerge from latent-variable models

By

Mia C. Morrell

Ilya M. Nemenman

Adviser

Audrey J. Sederberg

Co-Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Physics

2020

Acknowledgements

I thank Dr. Gordon Berman and Dr. Stefan Boettcher for forming the committee for this thesis.

I thank Dr. Audrey J. Sederberg for providing invaluable advice and mentorship.

I thank Alan C. W. Rohrbach for his love, support, advice, and expertise in computer construction and software distribution.

I thank my family for their love and their support of my academic endeavors.

Chiefly, I thank Dr. Ilya M. Nemenman for advising me throughout the course of this work's development and providing inspiration. Without his support and mentorship, this work would not exist.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Spatial and temporal criticality emerge from latent-variable models

Mia C. Morrell

Department of Physics, Emory University

Ilya Nemenman

Advisor

Audrey J. Sederberg

Co-advisor

(Dated: April 27, 2020)

I. INTRODUCTION

A. Motivation

In recent years high-throughput biological experiments have amassed increasingly large amounts of biological data. These vast data sets are high-dimensional and therefore are a challenge to interpret. A promising resolution to this problem is to adapt the Renormalization Group (RG) [2] framework for coarse-graining systems in statistical physics to extract relevant features and large-scale behaviors in biological data sets as well.

Recently, RG-inspired coarse-graining of large-scale neurophysiological data showed an intriguing emergence of nontrivial scaling behaviors in the activity of neural populations [1]. Specifically, the authors worked with 1000+ neurons in the mouse hippocampus as the animal repeatedly ran through a virtual maze. Their coarse-graining scheme involved combining the most correlated neurons into neural clusters by analogy with Kadanoff's hyperspins [3], using correlation as a proxy for locality. Various correlation functions of neural clusters exhibited self-similarity for different cluster sizes, indicative of criticality. Further analysis inspired by Wilson's momentum space approach [4] to renormalization surprisingly revealed that the joint distribution of activities of the clusters flowed to a non-Gaussian fixed point. The mechanisms responsible for these behaviors remain unknown, and thus it is unclear which other systems may exhibit them.

One observed signature of biological criticality is the Zipf's's law, which describes a power law relation between the rank and the frequency of a system's states [5]. It has been explained by the existence of stationary latent (unobserved) fields (such as stimuli or internal states) that couple neurons (spins) over long distances [6]. Similarly, in this Thesis, we show that the observations of Ref [1], including scaling properties of the free energy, the cluster-cluster correlations, the cluster autocorrelations, and the flow of the cluster activity distribution to a non-Gaussian fixed point can be explained, within experimental error, by a model of non-interacting neurons coupled to latent *dynamical* fields. This is the first model to explain such a variety of spatio-temporal phenomena under RG coarse-graining.

In this Thesis, we will first introduce the necessary experimental and theoretical background to understand our approach. We then introduce our model, perform the coarse-graining of Ref. [1] on data generated from it, and compare our findings with experimental results.

II. EXPERIMENTAL BACKGROUND

First, we will discuss the experimental setup of Ref. [1] and note important features of the rodent hippocampus we need to consider when writing down a model of this system.

A. Experimental setup

In this thesis, we consider the typical experimental setup used in studies of hippocampal neuronal networks in rodents, as used in Ref [1]. The rodent is placed on a rotating styrofoam ball which acts a treadmill and immersed in a virtual reality environment which provides a linear track that the rodent is allowed to run on. Once the rodent has reached the end of the virtual track, the program restarts and the rodent runs the track again. Thus, we can consider this setup as continuous runs on a toroidal track.

B. Place fields

Since in this Thesis we are considering a system of neurons in the mouse hippocampus, we must note experimentally found characteristics of this system. Spatial location is perhaps the most extensively researched of observed stimuli in the rodent hippocampus. Ref [1] notes the presence of cells called place cells: cells which help create a spatial map by which the rodent recognizes its own location.

Place cells are neurons in the hippocampus which tend to fire at a specific location [7]. The spatial neighborhood in which a given place cell becomes active is called its place field. In Ref [1], roughly half of the neurons recorded appear to be place cells.

Ref [8] uses the following procedure to classify hippocampal neurons into place and nonplace cells, a method also used in Ref [1]:

1. Bin the length of the track into spatial segments

2. Calculate the mean activity for each neuron within each of these bins

3. Calculate the spatial information for each neuron:

$$I = \sum_{k=0}^{K} p_k \frac{\langle s_i \rangle_k}{\langle s_i \rangle} \log_2 \frac{\langle s_i \rangle_k}{\langle s_i \rangle} \tag{1}$$

where there are K spatial bins spanning the track, and p_k is the probability of the rat being in the given bin k. Additionally, $\langle s_i \rangle$ is the overall average activity for cell i, and $\langle s_i \rangle_k$ is the average activity in spatial bin k for cell i.

4. Shuffle each neuron's activity recording 500 times, each time calculating the spatial information and saving it into a distribution

5. If the true spatial information of a given neuron falls at or above 80th percentile in the shuffled spatial information distribution for that neuron, then the given neuron is a place cell. If not, that neuron in a nonplace cell.

It is necessary to select a definitive approach to the phenomenology of place cells. There are two proposed mechanisms for place cells [8]:

1. Individual cells respond to spatial stimuli

2. Place cells interact and respond collectively to spatial stimuli in a network

In this thesis, we will assume the first and simplest mechanism.

III. THEORY

We now will discuss some theoretical background required to understand this Thesis. First, we will discuss why using spin models is an appropriate approach to modelling neural systems. Next, we will discuss the coarse-graining schemes which are used in the analysis of our simulations. We then provide a brief overview of critical phenomena relevant to this thesis, propose the scaling laws we might observe upon coarse-graining our simulations, and provide motivation on why we suspect that our simulations are close to a critical point.

A. Spin models applied to neural systems

In this section, we will address why spin models are appropriate for neural systems. In our rodent hippocampus system, it is natural to define two microstates for each cell we are recording from: active (firing) and inactive (not firing). Let there be N cells $s_n, n \in [1, N]$ which we record from. Let each s_n take on the value 1 if it is active, and 0 if it is inactive. Now, let us take a snapshot of this configuration of cells $\{s_n\}$ at discrete time steps over some time period.

Suppose we want to know the probability of observing any given cell configuration $\{s_n\}$ making as little assumptions as possible. We will derive such a probability distribution by using techniques pioneered by Ref [9].

We will use Shannon entropy to define the information entropy of this system:

$$S = -\sum_{k=1}^{2^N} p(\{s_i\}_k) \ln(p(\{s_i\}_k) \tag{2}$$

where k indexes all possible configurations $k \in [1, 2^N]$, and $p(\{s_i\}_k)$ denotes the probability of observing the configuration $\{s_i\}_k$.

In this system of N cells, a possible observable is the average activity $\langle s_n \rangle$ for each cell s_n. We will the set of observables $\{\langle s_n \rangle\}$ as constraints. We seek the probability distribution which maximizes Shannon entropy S while conforming to to the N constraints $\{\langle s_n \rangle\}$:

$$\langle s_n \rangle = \sum_{k=1}^{2^N} s_{i_k} p(\{s_i\}_k) \tag{3}$$

where s_{i_k} denotes the value of spin s_i in ensemble k. Additionally, we have a normalization constraint that the sum of the probabilities of observing each state is 1:

$$\sum_{k=1}^{2^N} p(\{s_i\}_k) = 1. \tag{4}$$

Using the typical maximization methods, we can set up the following loss function:

$$\mathscr{L} = -\sum_{k=1}^{2^N} p(\{s_i\}_k) \ln(p(\{s_i\}_k)) - \sum_{k=1}^{2^N} \sum_{i=1}^{N} h_i s_{i_k} p(\{s_i\}_k) - \alpha(\sum_{k=1}^{2^N} p(\{s_i\}_k) - 1) \tag{5}$$

where the set of $\{h_i\}$ and the scalar α are Lagrange multipliers. Now following the usual procedure, let $\frac{\partial \mathscr{L}}{\partial p(\{s_i\}_k)} = 0$ and solve for $p(\{s_i\}_k)$, normalizing by dividing the result by the previous normalization constraint. The result is the following:

$$p(\{s_i\}_k) = \frac{1}{Z} e^{-H(\{s_i\}_k} \tag{6}$$

which is a Boltzmann distribution using the following energy function:

$$H(\{s_i\}_k) = \sum_{i=1}^{N} h_i s_i \tag{7}$$

and the following partition function:

$$Z = \sum_{k=1}^{2^N} e^{-H(\{s_i\}_k)}. \tag{8}$$

This is an exact mapping to conditionally independent Ising model formalism. In this Thesis, we will apply this formalism with the following twist: the activity of a given cell is determined by the presence of multiple stimuli, or fields.

B. Coarse-graining schemes

In this Thesis, we hope to extract scaling laws in neural activity statistics by applying coarse-graining schemes. We use 2 distinct coarse-graining schemes: real-space and momentum-space. In this subsection, we will describe each of these coarse-graining schemes in detail.

1. Coarse-graining in real space

We replicate the real space coarse-graining analysis detailed in Ref [1]. A single real space coarse-graining iteration is accomplished in the following manner: (1) Calculate the correlation matrix C_{ij} for $\{s_i(t)\}$. (2) Select the rows and columns corresponding to the element $max(\{C_{ij}|i \neq j\}$ and set these to 0. (3) Add the activities of cells i and j, record in a new activity matrix. (4) Repeat this procedure until all elements of C_{ij} have been set to 0.

2. Coarse-graining in momentum space

We replicate the momentum space coarse-graining analysis detailed in Ref [1]. This is done by first calculating the covariance matrix Γ_{ij} of the neural activity matrix S_{it}, where i indexes neurons and t indexes time step. We then calculate the eigenvalues and eigenvectors of Γ_{ij} and construct

a matrix containing all of the eigenvectors in column vector form, ordered from the eigenvector with the largest corresponding eigenvalue to the eigenvector with the smallest corresponding eigenvalue. Since this matrix is just the fourier transform of S_{it}, we label it $\tilde{S}_{ij}$. Thus each column of $\tilde{S}_{ij}$ is a fourier mode of S_{it}.

We perform the coarse-graining by selecting the first k fourier modes and placing them in the matrix $\tilde{S'}_{ij'}$, where $j' \in [0,k)$. We then calculate the neural activity fluctuation matrix $\Phi_{it} = S_{it} - \langle S_{it} \rangle$ and compute

$$S_{it}^{(k)} = z_n \tilde{S}_{ij'} [\tilde{S}_{ij'}]^T \Phi_{it} \tag{9}$$

where z_n is a multiplier chosen such that

$$\langle [S_{it}^{(k)}]^2 \rangle = 1. \tag{10}$$

C. Critical Behavior and Scaling

Here we discuss the meaning of criticality: what exactly we hope to find by coarse-graining our simulations, and what it means for a system to be close to a critical point. First, we will discuss the definition of criticality in statistical mechanical systems. We will then discuss the Ornstein-Zernike correlation form, an important expression of the correlation function for critical systems.

1. A review of critical phenomena

This section is based on selected chapters from [2].

The phrase cmerritical phenomena denotes the behavior of thermodynamic observables in the neighborhood of a critical point , or a second-order phase transition. A second-order phase transition denotes a thermodynamic system in which its free energy and the first derivative of its free energy are continuous across a critical temperature T_C while the second derivative of its free energy is discontinuous. A common physical example of a critical point in a second-order phase transition is the liquid-gas transition. Another common example is the Curie point in a ferromagnetic system.

Consider a ferromagnetic system of spins with a Curie temperature T_C, and think about the

symmetries below and above T_C. Below T_C, the spins are magnetized and all want to point in one direction. Above T_C, the spins are not magnetized, and so the system is rotationally invariant. Thus when we decrease T from above T_C to below T_C, the rotational symmetry of our system is broken.

Thus we must have 2 different sets of functions for the thermodynamic observables, one for each phase: $(T > T_C)$ and $(T < T_C)$. These functions cannot be connected: that is, they cannot be extended continuously across T_C.

Here we encounter a problem. If we are close to the critical point T_C, how do we calculate the thermodynamic observables? The answer is to define an extra parameter called an order parameter. In the case of our ferromagnetic system, the order parameter will be the magnetization M. Note that the choice of which parameter to use as an order parameter for a system is often not obvious.

How can we quantify our usual thermodynamic observables close to T_C? First, we define the parameter $t = \frac{T-T_C}{T_C}$. Next, we assume that close to T_C, our thermodynamic observables may be expressed as a combination of a finite, or regular part and an infinite, or singular part. We assume that the singular part of each observable is some power law of t with a fractional exponent.

An important measure of systems at criticality is the correlation length ξ, which naturally arises from the Ornstein-Zernike form of the correlation function of an order parameter, as we shall see soon. The correlation length is a measure of spatial memory of the system, and diverges at T_C. The power laws for the singular part of our expressions for thermodynamic quantities, known as scaling laws , are typically deduced using dimensional analysis after assuming that a some quantity of dimension $[\text{length}]^{-D} \propto \xi^{-D}$ close to the critical point.

2. *The Ornstein-Zernike correlation form*

This section is based on selected chapters from [2].

The typical form of correlation used when discussing scaling laws of systems near criticality is the Ornstein-Zernike form of the correlation function. This form may be obtained by the following procedure:

Let some order parameter M be decomposed into some density function $m(\mathbf{r})$ such that

$$M = \langle \int d^3 r m(\mathbf{r}) \rangle. \tag{11}$$

Then let the usual correlation function as a function of the spatial vector $\mathbf{r}$ be

$$\Gamma(\mathbf{r}) = \langle m(\mathbf{r})m(0)\rangle - \langle m(\mathbf{r})\rangle\langle m(0)\rangle. \tag{12}$$

We assume that $\langle m(\mathbf{r})\rangle = \langle m(0)\rangle$, so:

$$\Gamma(\mathbf{r}) = \langle m(\mathbf{r})m(0)\rangle - \langle m(0)\rangle^2. \tag{13}$$

Then let $\tilde{m}(\mathbf{k}) = \int d^3 r e^{-i\mathbf{k}\cdot\mathbf{r}} m(\mathbf{r})$, the Fourier transform of m. Thus we have

$$m(\mathbf{k}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{m}(\mathbf{k}). \tag{14}$$

Now compute $\tilde{\Gamma}(\mathbf{k})$, the Fourier transform of $\Gamma(\mathbf{r})$:

$$\tilde{\Gamma}(\mathbf{k}) = \langle \tilde{m}(\mathbf{k})m(0)\rangle - \langle m(0)\rangle^2 (2\pi)^3 \delta(\mathbf{k}) \tag{15}$$

as the Fourier transform of $\langle m(0)\rangle^2$ is $\langle m(0)\rangle(2\pi)^3\delta(\mathbf{k})$, by the Fourier transform of a constant. Note that by definition, $\langle m(0)\rangle = 0$. Then we have:

$$\tilde{\Gamma}(\mathbf{k}) = \langle \tilde{m}(\mathbf{k})m(0)\rangle. \tag{16}$$

If we plug in $\mathbf{r} = 0$ to the previous integral, we obtain:

$$m(0) = \frac{1}{(2\pi)^3} \int d^3k \tilde{m}(\mathbf{k}). \tag{17}$$

We then import the general identity for any vectors $\mathbf{m}, \mathbf{p}$:

$$\langle \tilde{m}(\mathbf{k})\tilde{m}(\mathbf{p})\rangle = (2\pi)^3 \delta(\mathbf{k}+\mathbf{p}) |m(\mathbf{k})|^2. \tag{18}$$

Then we have:

$$\tilde{\Gamma}(\mathbf{k}) = \langle |\tilde{m}(\mathbf{k})|^2 \rangle \tag{19}$$

We now assume the free energy of this system is of them form:

$$\mathscr{G} = \int d^3r[c_1|\nabla m(\mathbf{r})|^2 + c_2 m^2(\mathbf{r})]. \tag{20}$$

Taking the Fourier transform, we obtain:

$$\tilde{\mathscr{G}} = \frac{d^3k}{(2\pi)^3}(c_1k^2 + c_2)|\tilde{m}(\mathbf{k}))|^2 \tag{21}$$

noting that the Fourier transform of a derivative such as $\nabla m(\mathbf{r}) = k\tilde{m}(\mathbf{k})$.

We then take the ensemble average of the kth Fourier mode of the free energy $\mathscr{G}$, obtaining:

$$\langle \mathscr{G}(\mathbf{k}) \rangle = kT \tag{22}$$

from equipartition of energy. Thus we have:

$$\langle |\tilde{m}(\mathbf{k})|^2 \rangle = \frac{kT}{c_1k^2 + c_2} = \tilde{\Gamma}(\mathbf{k}). \tag{23}$$

Then by an inverse Fourier transform, we receive:

$$\Gamma(\mathbf{r}) = \frac{1}{r}e^{-\frac{r}{\xi}} \tag{24}$$

where $\xi \equiv \sqrt{\frac{c_a}{c_2}}$ and is known as the correlation length. Of course, this derivation is done in only one dimension, $\mathbf{r}$. In d dimensions, the Ornsetein-Zernike correlation function is the following:

$$\Gamma(\mathbf{r}) \propto \frac{1}{r^p}e^{\frac{-r}{\xi}} \tag{25}$$

where $p = d - 2 + \eta$ where d denotes dimensionality.

Now when we approach a critical point, $\xi \to \infty$,

$$\Gamma(\mathbf{r}) \propto \frac{1}{r^p} \tag{26}$$

is the expression for the correlation function. We will use this function when predicting scaling laws for our simulations close to their critical points.

D. Predictions of scaling laws

In this section, we make predictions of scaling laws we might observe in our simulations if they are close to criticality. First, we discuss how activity variance might scale as a function of coarse-grained variable, or cluster, size. Next, we construct an effective free energy for our system and show how this free energy might scale with cluster size. We then discuss how the eigenvalue spectrum for each coarse-grained variable at each coarse-graining iteration displays scaling behavior with cluster size.

1. Scaling in activity variance

We can make predictions on the relation of the variance of the coarse-grained variable activity as a function of cluster size. If our system is at a critical point, then we expect the correlations between coarse-grained variables to have a self-similar structure. Thus, we expect our coarse-grained system to consist of variables which are somewhere between perfectly correlated and perfectly uncorrelated with one another. We claim that the variance of the coarse-grained variables will increase as a power law of cluster size with an exponent between 1 and 2.

We can prove this claim by calculating the upper and lower bounds of the variance of our system if it is close to criticality. We note that the upper bound for the coarse-grained activity variance in our system is that of perfectly correlated variables, while the lower bound is that of perfectly uncorrelated variables.

First, we will calculate the variance of perfectly uncorrelated, or independent variables. Note that for independent variables X and Y the covariant term $\langle XY\rangle - \langle X\rangle\langle Y\rangle$ is 0.

Proof. Suppose we have 2 independent variables, X and Y.

$Var(X+Y) = \langle (X+Y)^2\rangle - \langle X+Y\rangle^2 = \langle X^2\rangle + 2\langle XY\rangle + \langle Y^2\rangle - \langle X\rangle^2 - 2\langle X\rangle\langle Y\rangle - \langle Y\rangle^2 = (\langle X^2\rangle - \langle X\rangle^2) + (\langle Y^2\rangle - \langle Y\rangle^2) + 2(\langle XY\rangle - \langle X\rangle\langle Y\rangle)$

$\Longrightarrow$

$$Var(X) + Var(Y) = Var(X+Y) \tag{27}$$

Thus, the coarse-grained variance will increase in proportion to cluster size K. □

Next, we will calculate the variance of perfectly correlated variables. Note that for variables X

and Y, by Schwarz inequality the covariant term obeys:

$$|Cov(X,Y)|^2 = [\langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle]^2 \leq [\langle (X - \langle X \rangle)^2 (Y - \langle Y \rangle)^2 \rangle]. \tag{28}$$

Proof. Suppose variables X and Y are perfectly correlated. Thus their covariance Eq 28 is maximized such that:

$$|Cov(X,Y)|^2 = [\langle (X - \langle X \rangle)^2 (Y - \langle Y \rangle)^2 \rangle] = Var(X)Var(Y). \tag{29}$$

Thus:

$$|Cov(X,Y)| = \sqrt{Var(X)Var(Y)}. \tag{30}$$

Now we can plug in our covariant term into the usual equation of variance of the term $X + Y$:

$$Var(X+Y) = Var(X) + Var(Y) + 2Var(X)^{1/2} Var(Y)^{1/2}. \tag{31}$$

If our variables X and Y are perfectly correlated, $Var(X) = Var(Y)$. Then:

$$Var(X+Y) = 4Var(X). \tag{32}$$

Now consider that at every coarse-graining step, we are iterating the addition of 2 variables. If this procedure is performed with perfectly correlated variables at every coarse-graining step, the coarse-grained variance will increase in proportion to $4^k = 2^{2k} = K^2$, as claimed. □

Thus if our system is self-similar, we expect the coarse-grained activity variance to scale as a power of of cluster size K with some exponent between 1 and 2.

Now let's apply these findings to our simulated system of neurons which are being iteratively coarse-grained. In our system of neurons we can examine the activity variance over the coarse-grained variables such that

$$M_2(\{s_i^{(k)}\}) = \frac{K}{N} \sum_{i=1}^{N/K} [\langle (s_i^{(k)})^2 \rangle - \langle (s_i^{(k)}) \rangle^2] \tag{33}$$

at each coarse-graining iteration k. Now we again note that our coarse-grained variables are linear

combinations of our microscopic variables. Thus, if our microscopic variables were completely independent, we would expect $M_2(\{s_i^{(k)}\})$ to scale as

$$M_2(\{s_i^{(k)}\}) \propto K^{\alpha} \tag{34}$$

with $\alpha = 1$, as shown above. If our microscopic variables were fully correlated, we would expect $M_2(\{s_i^{(k)}\})$ to scale as Eq 34 with $\alpha = 2$, as shown above.

2. Free energy scaling

We can also make predictions concerning the distribution of neuron activity within coarse-grained variables when our system is close to a critical point. We will do this by constructing an effective free energy function and tracking how it evolves over coarse-graining iterations. We write down the probability that every member of a set of neurons will be silent:

$$P(\{s_i^{(0)}\}) = \frac{1}{Z} e^{-E(\{s_i^{(0)}\})} \tag{35}$$

where the energy

$$E(\{s_i^{(0)}\}) = \sum_{i=1}^{K} h_i s_i^{(0)} + \sum_{i,j=1}^{K} J_{ij} s_i^{(0)} s_j^{(0)} + ... \tag{36}$$

is a expansion in terms of inter-neuronal interactions. When the set of neurons $\{s_i^{(0)}\}$ is completely silent,

$$P_{silence}(\{s_i^{(0)}\}) = \frac{1}{Z} \tag{37}$$

and thus we can define

$$F(\{s_i^{(0)}\}) = -\ln P_{silence}(\{s_i^{(0)}\}) \tag{38}$$

where $F(\{s_i^{(0)}\})$ is effective free energy.
It is reasonable to expect that $F(\{s_i^{(0)}\})$ scales as a function of our length scale K such that

$$F(\{s_i^{(0)}\}) \propto K^{\tilde{\beta}} \tag{39}$$

close to criticality.

3. Scaling in eigenvalue spectra

To deduce how a critical system's eigenvalue spectra of coarse-grained variables behaves as a function of rank under repeated coarse-graining iterations, we note that the Fourier modes of a correlation matrix of variables with symmetric interactions are the eigenvectors of that matrix [1] [10].

We calculate the covariance matrix

$$C_{ij}^{(k)} \equiv \langle s_i^{(0)} s_j^{(0)} \rangle - \langle s_i^{(0)} \rangle \langle s_i^{(0)} \rangle \tag{40}$$

where (k) indexes coarse-graining iteration. We calculate Eq 40 for each coarse-grained variable and obtain its Fourier transform $\tilde{C}_{ij}^{(k)}$ by calculating its eigenvectors. We apply Eq 25, the Ornstein-Zernike form for correlation close to a critical point, yielding a power law decay in Fourier transformed correlation as a function of Fourier wavevector $\mathbf{k}$:

$$\Gamma(\mathbf{k}) \propto \frac{1}{|\mathbf{k}|^{2-\eta}}. \tag{41}$$

If we consider $|\mathbf{k}|^d$ as a volume for the eigenvector $\mathbf{k}$, we obtain rank $R \approx |\mathbf{k}|^d$ for the eigenvalue corresponding to $\mathbf{k}$. Thus,

$$\lambda_R \propto \left(\frac{K}{R}\right)^{\mu} \tag{42}$$

where $\mu \equiv \frac{2-\eta}{d}$ and K is cluster size. We note that the largest wavevector $\mathbf{k}$ should represent the highest frequency and thus have the smallest eigenvalue, and therefore highest rank R. Similarly, the smallest wavevector should represent the lowest frequency and thus should have the largest eigenvalue, and therefore the lowest rank R. This observation is consistent with Eq 42.

4. Scaling in correlation time

Yet another reasonable prediction we can make concerning our system is that close to criticality, correlation time τ_c is some power law of length scale with some exponent $\tilde{z}$ according to the dynamic scaling hypothesis. Now in our system of coarse–grained neurons, we have dynamics at a length scale of K, or cluster size. Thus we propose that correlation time τ_c is a power law of

cluster size K:

$$\tau_C \propto K^{\tilde{z}} \tag{43}$$

where $\tilde{z}$ is the dynamical critical exponent.

5. *Flow to a non-gaussian fixed point*

We can also predict how our system's distribution of coarse-grained variables evolves over coarse-graining scheme iterations when it is close to criticality. Consider our coarse-graining scheme in momentum space. Note that the coarse-grained variables are linear combinations of the original variables. As long as the correlations between original variables are sufficiently weak, their distribution will approach a Gaussian distribution due to the central limit theorem. However, close to criticality the system may approach a non-Gaussian fixed point. We propose that we will find such an approach to a non-Gaussian fixed point in our simulations.

E. Theoretical Motivation

In this section, we will discuss why we expect critical behavior in our simulations due to our choice of a latent variable model. First, we will discuss Zipf's Law and its connections to critical behavior and show that large systems coupled to latent fields may show critical behavior under certain conditions. Next, we will discuss how our model differs from other latent variable models such as that used in Ref [6].

1. *Zipf's Law*

The Zipf's Law is a phenomenon which has been shown to be intimately connected with large systems at criticality [6]. The Zipf's Law states that the rank of a state according to frequency is inversely proportional to the state's frequency. That is:

$$P(\{s_i\}_k) \propto (R(\{s_i\}_k))^{-\alpha}. \tag{44}$$

Specifically, the proportionality constant is the inverse of the Riemann zeta function evaluated at α, $\frac{1}{\zeta(\alpha)}$. Zipf's Law behavior has been demonstrated in a variety of complex systems such as word

frequency in language.

2. Criticality emerges naturally from large systems obeying Zipf's Law

Ref [5] found that critical behavior is a natural feature of large systems displaying Zipf's law behavior. Here I reproduce an argument made by Ref [5] in a bit more detail.

Claim: *Entropy is an exactly linear function of energy for large system size.*

Proof. Import the definition of Boltzmann entropy

$$S(H) = k_B \ln \Omega(H). \tag{45}$$

Then

$$S(H) = k_B \ln(\rho_{\delta H}(H)\delta H) = \ln(\rho_{\delta H}(H)) + \ln(\delta H) = \ln(\rho_{\delta H}(H)) + s_1 \tag{46}$$

for $k_B = 1$ and constant δH. Note that setting k_B to 1 is equivalent to using the Shannon information entropy in its Boltzmann-like form, and s_1 is subextensive, meaning it decreases slower than proportional to system size. Note also that entropy and energy are both extensive properties which do increase proportionally with system size N.

Thus

$$\ln(\rho_{\delta H}) = S(H) - s_1. \tag{47}$$

Note that $\rho_{\delta H}$ is the density of states such that $\rho_{\delta H} = \frac{1}{\delta H}\sum_k \mathbb{I}[H < H(\{s_i\}_k) < H + \delta H]$, where $\mathbb{I}$ is an indicator function which produces 1 for each H in range $H < H(\{s_i\}_k) < H + \delta H$, and 0 for all other H.

We now consider the cumulative density of states

$$\mathcal{N}(H) = \sum_k \mathbb{I}[H(\{s_i\}_k) < H] = \int_{-\infty}^{H} dH' \rho_{\delta H \to 0}. \tag{48}$$

Now we know that the density of states is as follows:

$$\rho_{\delta H} = e^{S(H) - s_1} = e^{N(s(\varepsilon \equiv \frac{H}{N})) - s_1} \tag{49}$$

where

$$s(\varepsilon) \equiv \lim_{N\to\infty} \frac{1}{N} S(\varepsilon \equiv \frac{H}{N}) \tag{50}$$

is a entropy density function such that

$$\int_{\varepsilon'} d\varepsilon' s(\varepsilon') = Ns(\varepsilon) = S(H). \tag{51}$$

Then

$$\mathscr{N}(H) = \int_{-\infty}^{H} dH' \rho_{\delta H\to 0} = \int_{-\infty}^{H} dH' e^{S(H')-s_1} = \int_{-\infty}^{H} dH' e^{Ns(\frac{H'}{N})+s_1} = N\int_{-\infty}^{H/N} d\varepsilon' e^{N(s(\varepsilon')+\frac{s_1}{N})} \tag{52}$$

with $\varepsilon' \equiv \frac{H'}{N}$. Now for large N, we may evaluate the integral:

$$\mathscr{N}(H) \approx e^{Ns(\frac{H}{N})} \to S(H) = \ln \mathscr{N}. \tag{53}$$

Now recall that the rank $R(\{s_i\}_k)$ of a state $\{s_i\}_k$ is exactly equivalent to $\mathscr{N}(H(\{s_i\}_k = H)$. Thus we have

$$S(H) = \ln(R(\{s_i\}_k)) \tag{54}$$

Now import the Zipf's law:

$$P(\{s_i\}_k) = \frac{1}{\zeta(\alpha)} (R(\{s_i\}_k))^{-\alpha} \tag{55}$$

and take its logarithm:

$$\ln(P(\{s_i\}_k)) = -ln(\zeta(\alpha)) - \alpha \ln(R(\{s_i\}_k)). \tag{56}$$

Finally, import the following:

$$P(\{s_i\}_k) = \frac{1}{Z} e^{\frac{-H(\{s_i\}_k)}{k_B T}} \tag{57}$$

where we have already set k_B to 1, and proceed to set T and Z to 1 without loss of generality. Thus we have

$$\ln(P(\{s_i\}_k)) = -H(\{s_i\}_k). \tag{58}$$

Additionally, import our previously derived expression $S(H(\{s_i\}_k)) = \ln(R(\{s_i\}_k))$.

Plug these into our expression $\ln(P(\{s_i\}_k)) = -\ln(\zeta(\alpha)) - \alpha \ln(R(\{s_i\}_k))$ and we receive the

following relation:

$$H = \ln(\zeta(\alpha)) + \alpha S(H). \tag{59}$$

But $\ln(\zeta(\alpha))$ is subextensive, so for large system size N:

$$S(H) = \frac{1}{\alpha} H + \dots . \tag{60}$$

Now it is clear that for all energies H and sufficiently large N, the following is true:

$$\frac{d^2 S}{dH^2} = 0 \tag{61}$$

a strong indication of a critical point for large system size N displaying Zipf's law behavior.

□

To summarize, we have learned that critical behavior is a natural feature arising from large systems displaying the Zipf's law. This phenomenon is referred to as Zipf's criticality.

3. Large systems coupled to latent stimuli obey Zipf's Law

Building on this emergence of critical behavior from large systems displaying the Zipf's law, Ref [6] demonstrated both analytically and numerically that large systems coupled to hidden stimuli display Zipf's criticality. The implication of this finding is that systems which display Zipf's criticality, such as real neuronal systems, may do so because they are coupled to many unknown stimuli.

Here I present the analytical argument made by Ref [6]:
Suppose we have a set of N conditionally independent spins $\{s_i\}$ in which each $s_i \in \{-1, 1\}$. The set $\{s_i\}$ is coupled to a hidden stimulus h. h is drawn from some distribution $\mathscr{D}(h)$, and varies rapidly with time. Now we have the following probability of observing the configuration of spins $\{s_i\}_k$ given h:

$$P(\{s_i\}_k|h) = \prod_{i=1}^{N} P(\{s_i\}_k|h) = \prod_{i=1}^{N} \frac{e^{hs_i}}{2\cosh h}. \tag{62}$$

And thus the probability of observing the configuration of spins $\{s_i\}_k$ is the previous equation

marginalized over h:

$$P(\{s_i\}_k) = P(m) = \frac{1}{2^N}\int dh \mathscr{D}(h) e^{N(hm - \ln\cosh h)} = \frac{1}{2^N}\int dh \mathscr{D}(h) e^{N\mathscr{H}(m,h)} \tag{63}$$

where $m = \sum_i \frac{s_i}{N}$ is the average magnetization of the spin system, and we define $\mathscr{H}(m,h) \equiv hm - \ln\cosh h$.

We now perform a saddle-point approximation on the previous integral and procure:

$$P(\{s_i\}_k) \approx 2^{-N}\mathscr{D}(h^*) e^{N(h^* m - \ln\cosh h^*)} \tag{64}$$

where h^* is the maximum likelihood estimate of h given $\{s_i\}_k$. Note that for us to perform this saddle point approximation, the Fisher information must obey $\mathscr{F}(h^*) \equiv -N\frac{\partial^2 \mathscr{H}}{\partial h^2}|_{h^*} = N(1+m^2) >> 1$. To obtain this expression we use the identity $1 + tan^2 x = sec^2 x$ and the result from the saddle point approximation: $m = \tanh h^*$.

We now use the following identities:

$$\tanh^{-1} m = \frac{1}{2}\ln(\frac{1+m}{1-m}) \tag{65}$$

and

$$\cosh[\tanh^{-1} m] = (1-m^2)^{-\frac{1}{2}} \tag{66}$$

and substitute our probability expression $P(\{s_i\}_k) \approx 2^{-N}\mathscr{D}(h^*) e^{N(h^* m - \ln\cosh h^*)}$ into the previously used expression relating energy of a given state and log probability of that state $-\ln(P(\{s_i\}_k)) = H(\{s_i\}_k)$.

The result is the following, recalling that $h^* = \tanh^{-1} m$:

$$\begin{aligned}
H(m) &\approx -N(h^* m - \ln\cosh h^*)\ln(2^{-N}\mathscr{D}(h^*)) \\
&= -N(\frac{1}{2}\ln(\frac{1+m}{1-m})m - \ln(1-m^2)^{-\frac{1}{2}})\ln(2^{-N}\mathscr{D}(h^*)) \\
&= -N(\frac{1}{2}(m\ln(1+m) - m\ln(1-m)) + \frac{1}{2}(1+m) + \frac{1}{2}\ln(1-m))\ln(2^{-N}\mathscr{D}(h^*)) \\
&= -N(\frac{1}{2}(1+m)\ln(1+m) + \frac{1}{2}(1-m)\ln(1-m))(\ln(2^{-N})) \\
&= -N(\frac{1}{2}(1+m)\ln(\frac{1+m}{2}) + \frac{1}{2}(1-m)\ln(\frac{1-m}{2})) \\
&= N\varepsilon(m)
\end{aligned}$$

where we neglect the $\mathscr{D}(h^*)$ within the natural log and define $\varepsilon(m) \equiv \frac{1}{2}(1+m)\ln(\frac{1+m}{2}) + \frac{1}{2}(1-m)\ln(\frac{1-m}{2})$.

Now we can connect the expression $N\varepsilon(m)$ directly to the entropy of the system as a function of average magnitization, $S(m)$. In our system of N conditionally independent spins, there must be $K(m) = N(\frac{1+m}{2})$ spin up states for a given average magnetization m. Then there are $\binom{N}{K}$ states where K spins are spin up. Then we calculate the entropy of these configurations: $S(m) = \ln\binom{N}{K} = \ln(\frac{N!}{K!(N-K)!}) = \ln N! - \ln K! - \ln(N-K)!$ where $(N-K) = (\frac{N(1-m)}{2})$. Then using Stirling's approximation, we can crunch through some algebra and determine the following:

$$S(m) = -N(\frac{1}{2}(1+m)\ln(\frac{1+m}{2}) + \frac{1}{2}(1-m)\ln(\frac{1-m}{2})) = N\varepsilon(m). \tag{67}$$

Thus we have shown the following is true for our system of conditionally independent spins:

$$S(m) = H(m) \tag{68}$$

for sufficiently large system size N. Ref [6] goes on to verify that this equivalence of entropy and energy for given average magnetization m holds for finite system size N through numerical simulations.

The importance of this result cannot be understated. We have shown in a large system of conditionally independent spins coupled to a hidden varying stimulus, Zipf's criticality arises a natural feature of the system!

4. *Requirements for our model*

In this subsection, we have discussed Zipf's Law and its connections to critical behavior and shown that large systems coupled to latent fields may show critical behavior, a characteristic which motivates our choice of a latent variable model to explain the results of Ref [1]. Here we note a few differences between the latent variable model presented in Ref [6] and our latent variable model. First, the latent variable model used in Ref [6] involved stationary latent fields, while we seek to apply dynamic latent fields in order to replicate the dynamic hidden stimuli which influence the rodent hippocampus. Second, we will include known, periodic stimuli in our model to replicate experimentally found features of the rodent hippocampus [8].

By constructing a latent variable model according to these specifications, we expect to construct simulated neural activity which displays critical behavior upon application of the coarse-graining schemes detailed previously in this section. If such a feat can be accomplished, perhaps the influence of hidden stimuli on the rodent hippocampus explains the emergence of critical behavior as found in Ref [1].

IV. METHODS

In the previous sections, we discussed the experimental and theoretical basis for this Thesis, predicted scaling laws in thermodynamic observables we might observe upon coarse-graining our simulations using the coarse-graining schemes detailed in *Theory: Coarse-graining schemes*, and explained why we expect these scaling laws to emerge using reasoning inspired by Ref. [6]. In this section, we present a latent variable model and show that the scaling laws observed by Ref. [1] and presented in *Theory: Predictions of scaling laws* can be generated from this model, often within experimental error.

A. The model

Here we will introduce the latent variable model that generates our simulated data. We consider N binary neurons (spins) $s_i \in \{0,1\}, i \in [1,N]$, where $s_i = 0$ or 1 corresponds to a neuron being silent or active. The neurons are conditionally independent and coupled only by n_f latent fields

$h_m(t)$, $m \in [1, n_f]$ such that the probability of a population being in a certain state is

$$P(\{s_i\}|\{h_m\}) = \frac{1}{Z(\{h_m\})} e^{-H(\{s_i\},\{h_m\})}, \tag{69}$$

where Z is a normalization factor, and H is the "energy":

$$H = \eta \left(\sum_{i,m}^{N,n_f} h_m(t) W_{m,i} s_i + \varepsilon s_i \right). \tag{70}$$

Here ε stands for the bias toward silence, η controls the variance of individual neuron activity, and $W_{m,i}$ are the coupling constants that define which fields drive which neurons. While other choices are possible, we start with $N = 1024$ as in Ref. [1] and $n_f = 10$ (see *Discussion* for other choices).

The mouse hippocampus has neurons that are activated by the animal being near a certain point on the virtual track, which are called place cells. To reproduce experimental results within error bars, we need to account for the existence of place cells, even though they are not crucial for the emergence of scaling (see *Discussion*). Therefore, there are two different types of couplings $W_{m,i}$: the couplings to spatial fields and the couplings to latent fields.

The couplings to spatial fields

$$W_{m,i}^{(place)} = \begin{cases} 0 & \text{if } i \text{ is not a place cell} \\ \Gamma(1,1) & \text{if } i \text{ is a place cell} \end{cases} \tag{71}$$

correspond to the place fields $h_m^{(place)}(x)$, which are Gaussian waveforms with center $\nu_m \in (0, X]$ and standard deviation σ_m drawn from a gamma distribution with mean $X/10$ and standard deviation $X/20$, where X represents track length. We couple a neuron to a place field with constant probability r.

A Gaussian is an appropriate model of a place field, which drives a place cell to fire in the neighborhood of a certain location in space. We define the mouse location as $x \in (0, X]$ where $x = vt$ and v is the mouse velocity. To account for the many track runs recorded per experiment [1], mouse location x is reset to 0 after every track loop is completed. Each simulation consists of 200 of these track loops, and runtime is on the order of seconds on a laptop computer.

TABLE I. Default parameters. Parameter sweeps are performed by varying each of these parameters, one at a time.

Default parameter values		
Parameter	Description	Value
ϕ	latent field multiplier	$\phi = 1.0$
ε	bias towards silence	$\varepsilon = -16./6.$
η	variance multiplier	$\eta = 6.0$
q	probability of coupling to latent field	$q = 1.0$
n_f	number of latent fields	$n_f = 10$
τ	latent field time constant	$\tau = 0.1$
$h_m^{(place)}$	presence or absence of place fields	all cells coupled to latent fields, $N/2$ cells coupled to place fields

The couplings to the latent, or nonplace fields

$$W_{m,i}^{(latent)} = \phi \times \begin{cases} 0 & \text{if } i \text{ is not a nonplace cell} \\ \mathcal{N}(0,1) & \text{if } i \text{ is a nonplace cell} \end{cases} \tag{72}$$

correspond to the latent fields $h_m^{(latent)}(t)$, which are Orstein-Uhlenbeck processes of mean 0, unit variance, and time constant τ_m. As tabulated in TABLE I, we present results in FIG 1 through FIG 6 with all latent fields $h_m^{(latent)}$ possessing the same time constant τ_m. We couple a neuron to a latent field with constant probability q.

While we explored many different choices of parameters (see TABLE II), here we present results with $N = 1024$ as in Ref. [1], and $n_f = 10$. In TABLE I, we tabulate the parameter choices which produce the figures shown in this Thesis. We choose 50% of neurons to be place neurons, each coupled to its own place field (ν_m, σ_m), which is consistent with Ref. [1]. Each latent field is coupled to every neuron. Thus in our typical simulations, about 512 neurons respond to place and latent stimuli, and about 512 are exclusively latent-stimuli neurons.

V. RESULTS

In the following section, we replicate the analyses performed by Ref. [1] line by line. We perform simulations according to Eq. 69, follow the real space and momentum space coarse-

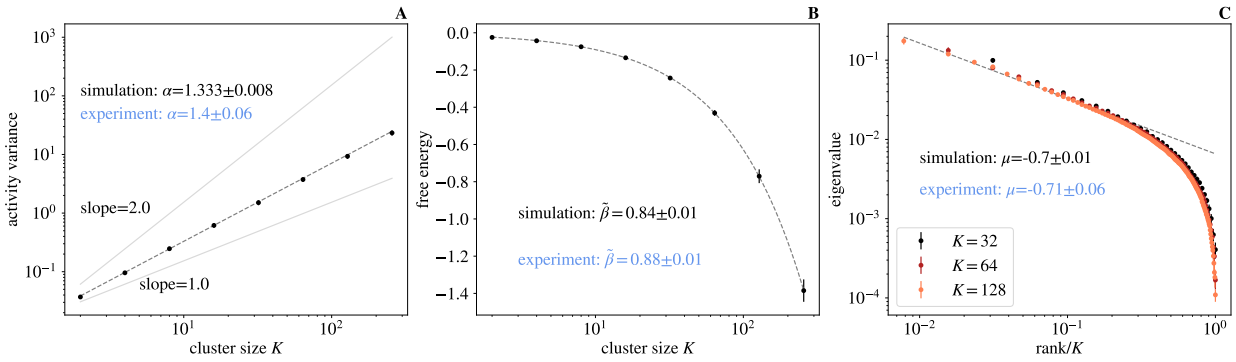

FIG. 1. **(A)** Activity variance over coarse-grained variables at each coarse-graining iteration. Fit to $M_2(\{s_i^{(k)}\}) \propto K^{\alpha}$ with $\alpha = 1.36 \pm 0.01$, where $s_i^{(k)}$ represents a coarse-grained variable at coarse-graining iteration k and K indicates cluster size. This is within error of the experimental observation $\alpha = 1.4 \pm 0.06$ (Ref. [1]), shown in blue. **(B)** Average free energy (Eq. 73) at each coarse-graining iteration. We observe that the average free energy at each coarse-graining iteration closely follows $F \propto K^{\tilde{\beta}}$ with $\tilde{\beta} = 0.82 \pm 0.01$. This is close to the experimentally found $\tilde{\beta} = 0.88 \pm 0.01$ (Ref. [1]), shown in blue. **(C)** Average eigenvalue spectrum of cluster covariance for cluster sizes $K = 32, 64, 128$. We observe behavior obeying Eq. 74 for roughly 1.5 decades with $\mu = 0.73 \pm 0.02$. This is within error of the experimental $\mu = 0.71 \pm 0.06$ (Ref. [1]), shown in blue. For **(A)**, **(B)**, **(C)**, error bars are standard deviations over randomly selected contiguous quarters of the simulation.

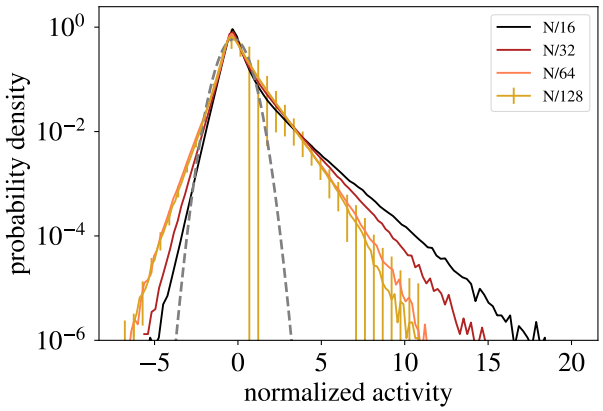

FIG. 2. Distribution of coarse-grained variables for $k = N/16, N/32, N/64, N/128$ modes retained. The distribution of coarse-grained variables approaches a non-gaussian distribution as k decreases. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

graining schemes of Ref. [1], and record the activity distributions of variables within clusters as we iterate the coarse-graining algorithms. See *Theory: Coarse-graining schemes* for more details on coarse-graining schemes. We note that all results produced by Ref. [1] can be quantitatively reproduced by our model, and we include corresponding experimental results in blue on each figure when appropriate. See *Theory: Predictions of scaling laws* for more details on the scaling

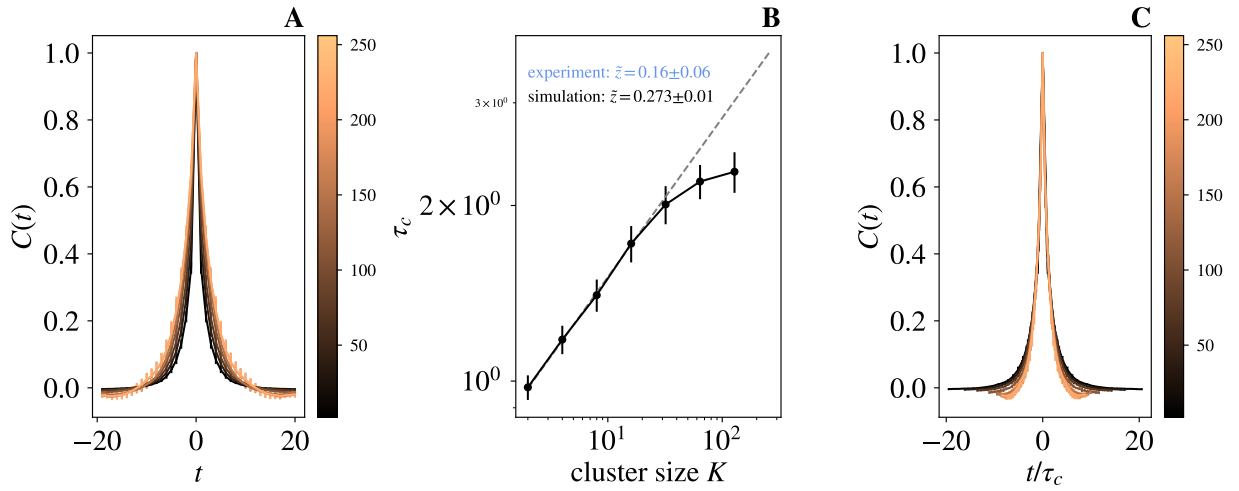

FIG. 3. **(A)** Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$ as a function of time. Colorbar indicates cluster size. **(B)** Time constants τ_c extracted from each curve in **(A)**, and observe behavior obeying $\tau_c \propto K^{\tilde{z}}$ for roughly 1 decade. We find exponent $\tilde{z} = 0.27 \pm 0.01$. This is close to the experimentally found $\tilde{z} = 0.16 \pm 0.06$ (Ref. [1]). **(C)** Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$, where time is rescaled by the appropriate τ_c for that coarse-graining iteration. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

laws observed by Ref. [1].

1. *Scaling of activity variance*

We can examine the activity variance over the coarse-grained variables at each coarse-graining iteration k. Now we again note that our coarse-grained variables are linear combinations of our microscopic variables. If our original variables were independent, we would expect the activity variance over coarse-grained variables to scale as a power law of cluster size K with an exponent of 1. If our original variables were maximally correlated, we would expect the activity variance over coarse-grained variables to scale as a power law of cluster size K with an exponent of 2. See *Predictions of scaling laws: Scaling in activity variance* for a derivation of these claims.

In Fig. 1A, we observe that coarse-grained activity variance is a power law of cluster size K with exponent 1.36 ± 0.01, suggesting that our simulation maintains a self-similar structure throughout the coarse-graining procedure. This result is within error bars of the experimental results.

2. *Scaling in free energy.* — We consider the distribution of neuron activity within coarse-grained variables. Following the procedure detailed in *Predictions of scaling laws: Scaling in free*

energy, we can define

$$F(\{s_i\}) = -\ln P(\{s_i = 0\}|\{h_m\}) \tag{73}$$

where $F(\{s_i\})$ is effective free energy. In Fig. 1B we observe that the average free energy at each coarse-graining iteration closely follows this prediction, producing a scaling exponent within error bars of experimental results [1].

2. *Scaling in eigenvalue spectra*

We examine the eigenvalue spectra of the coarse-grained variables at every coarse-graining iteration. if our simulations are close to criticality, the eigenvalue spectra of the coarse-grained variables at every coarse-graining step should obey the following scaling law which was proposed in *Predictions of scaling laws: Scaling in eigenvalue spectra*:

$$\lambda_R \propto \left(\frac{K}{R}\right)^{\mu} \tag{74}$$

where $\mu \propto \frac{1}{d}$ and K is cluster size and d is dimensionality.

In Fig. 1D we plot the average eigenvalue spectrum of the covariance matrix for each coarse-grained variable for cluster sizes $K = 32, 64, 128$. We observe behavior obeying Eq 74 for roughly 1.5 decades, suggesting that our simulation is close to a critical point. This result is within error bars of the experimental results [1]. Additionally, the collapse featured in Fig. 1D demonstrates that the average eigenvalue spectrum per coarse-graining iteration varies only as a function of K/R, an additional scale-invariant feature.

3. *Scaling in correlation time*

In *Predictions of scaling laws: Scaling in correlation time* we predicted that correlation time τ_c is some power law of length scale with some exponent $\tilde{z}$ according to the dynamic scaling hypothesis if our system is close to a critical point.

In Fig. 3A we plot the average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$. We observe a power law relation between the average autocorrelation function and cluster size K for roughly 1 decade in Fig. 3B with a scaling exponent within error bars of experimental results. Additionally, the collapse shown in Fig. 1C suggests we have found the appropriate time constants

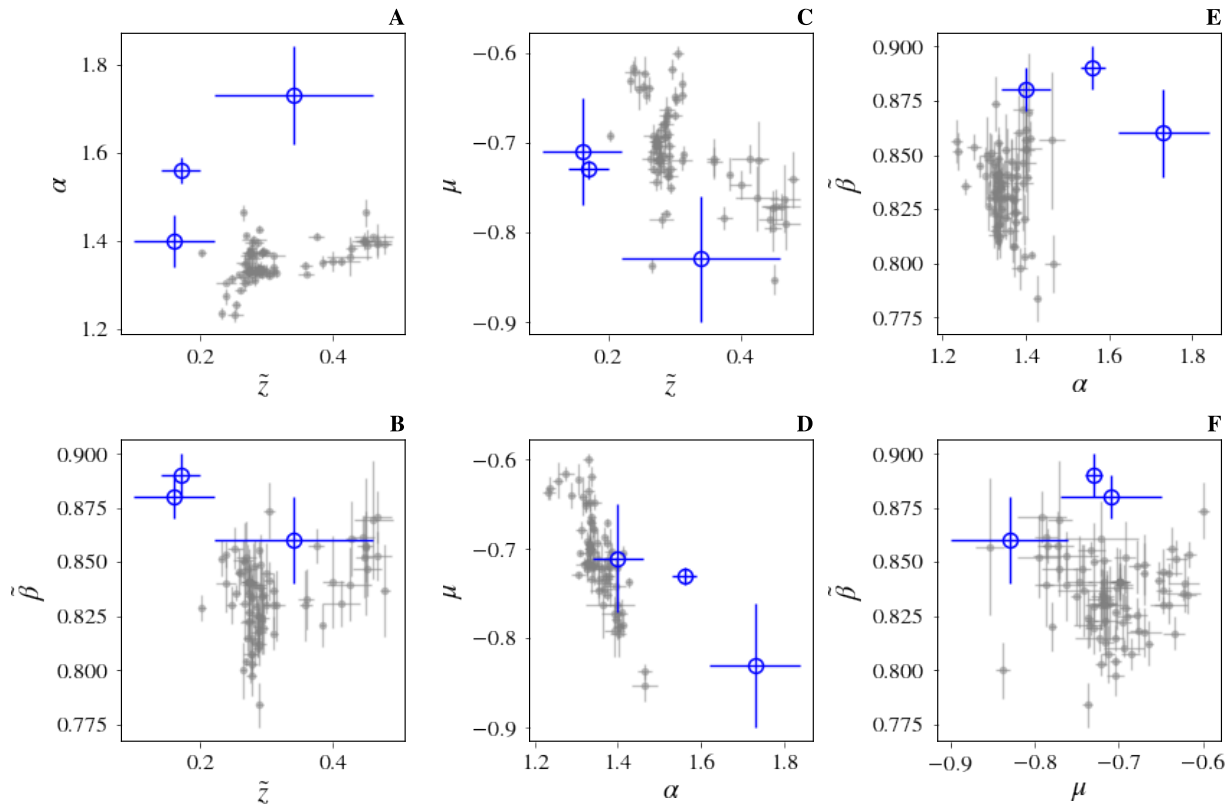

FIG. 4. **(A), (B), (C), (D), (E), (F)** Scaling exponents $\alpha, \tilde{\beta}, \mu, \tilde{z}$ generated by simulations with varying parameters against one another in gray. Experimental results generated by [1] in blue. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

τ_c which ensure $C(t/\tau_c)$ is scale invariant. This scaling behavior suggests that our simulations are indeed close the experimentally observed critical behavior [1].

4. *Flow to a non-gaussian fixed point*

In *Predictions of scaling laws: Flow to a non-gaussian fixed point* we predicted that the distribution of coarse-grained variables will approach a fixed, non-Gaussian distribution as we iterate the coarse-graining scheme.

In Fig. 2 we show the distribution of coarse-grained variables for $k = N/16, N/32, N/64, N/128$ modes retained. We observe that the distribution of coarse-grained variables approaches a non-Gaussian distribution as k decreases, suggesting that our simulation is close to a critical point.

TABLE II. Results from parameter sweeps over η, ϕ, and ε in Eq. 70, the latent field correlation time τ, the number of latent fields n_f, and the probability that a neuron couples to a latent field q. We vary these parameters one at a time while keeping all other parameters constant at default values which produce critical behavior. In the 3rd column, we record in which parameter regimes the simulation is critical.

Parameter sweeps			
Parameter	Description	Critical values	Comments
ϕ	latent field multiplier, sweep over $\phi = 0.8$ to $\phi = 1.5$	$\phi \in [0.9, 1.2]$	Latent fields which are too strong destroy the presence of place cells over coarse-graining iterations, and latent fields which are too weak are deleterious to variance scaling (FIG 28).
ε	bias towards silence, sweep over $\varepsilon = -1$ to $\varepsilon = -20$	$\varepsilon < -1.92$	Damaged eigenvalue scaling for $\varepsilon > -1.92$ (FIG 67). Coarse-grained distributions of activity from simulations with $\varepsilon > -1.5$ approach a gaussian fixed point (FIG 73).
η	variance multiplier, sweep over $\eta = 2.8$ to $\eta = 6.6$	$\eta \in [2.8, 6.6]$	Varying η does not impact scaling behavior (FIG 74 – FIG 83).
q	probability of coupling to latent field, sweep over $q = 0.25$ to $q = 1.0$	$q \geq 0.5$	No significant impact on scaling for $q \geq 0.5$, but $q < 0.5$ is deleterious to variance scaling (FIG 58).
n_f	number of latent fields	$n_f \geq 5$	5 or more latent fields are required to observe scaling behavior (FIG 37).
τ	latent field time constant, sweep over $\tau = 0.05$ to $\tau = 1.2$	$\tau \in [0.05, 1.2]$	Varying τ has no significant impact on scaling (FIG 44– FIG 53).
$h_m^{(place)}$	presence or absence of place fields	latent fields present	Simulations coupled only to place fields do not show scaling behavior (FIG 8 – FIG 14), and simulations with $N/2$ cells coupled only to place fields and $N/2$ cells coupled only to latent fields show a damaged variance scaling relation (FIG 10).

5. *Experimental agreement*

To compare our simulations to the experimental results generated in [1], we record the scaling exponents $\alpha, \tilde{\beta}, \mu, \tilde{z}$ of simulations with eigenvalue spectra collapse for at least 1.5 decades as shown in FIG.1D and activity variance scaling for over 2 decades as shown in FIG.1A. We plot these exponents against one another in pairs (FIG. 4), and plot the experimental results generated by Ref. [1] in blue.

We note that the exponents α and $\tilde{z}$ were less well-matched than the exponents μ and $\tilde{\beta}$. We expect that simulations involving latent fields with non-uniform time constants may produce improved results for the dynamic scaling exponent $\tilde{z}$. Nevertheless, we observe that the simulation

results shown in FIG. 4 align well with experimental results from Ref [1].

A. Discussion

To investigate which parameter regimes may give rise to criticality in our model, we vary the parameters η, ϕ, and ε in Eq. 70, the latent field correlation time τ, the number of latent fields n_f, and the probability that a neuron couples to a latent field q. We vary these parameters one at a time while keeping all other parameters constant at default values which produce critical behavior. We also run a simulation with only nonplace fields $h_m^{(latent)}$ included. We record parameters whose simulations display eigenvalue spectra collapse for at least 1.5 decades as shown in FIG. 1D and activity variance scaling for over 2 decades as shown in FIG. 1A. We summarize the parameter regime in which criticality arises in TABLE II.

By coarse-graining our simulations of binary neurons in the style of Ref. [1] and extracting scaling exponents within experimental range for eigenvalue spectra, activity variance, free energy, and correlation time, we have shown that the results of Ref. [1] are not unique to neural systems and are a result of the presence of unknown, time-varying stimuli. In fact, we expect the signatures of self-similarity discussed in this Thesis to emerge from any sparsely active system driven by 5 or more latent processes. See *Parameter sweeps* for more information regarding this claim.

VI. SUPPLEMENTARY INFORMATION

A. Supplementary figures

We also note that our simulation can reconstruct the place cell activity demonstrated in Ref. [1]. In FIG. 6 we show that definite place cells exist in our simulations, and that these place cells are robust over many coarse-graining iterations.

In FIG. 5A we plot the pairwise correlation coefficients of our simulated neurons. We note that FIG. 5A is qualitatively similar to the analogous result produced in Ref. [1]. Features reconstructed from Ref. [1] include a sharp peak in density just to the right of 0, a short left tail, and a long right tail in pairwise correlation coefficient distribution ending at correlation coefficients greater than 0.6.

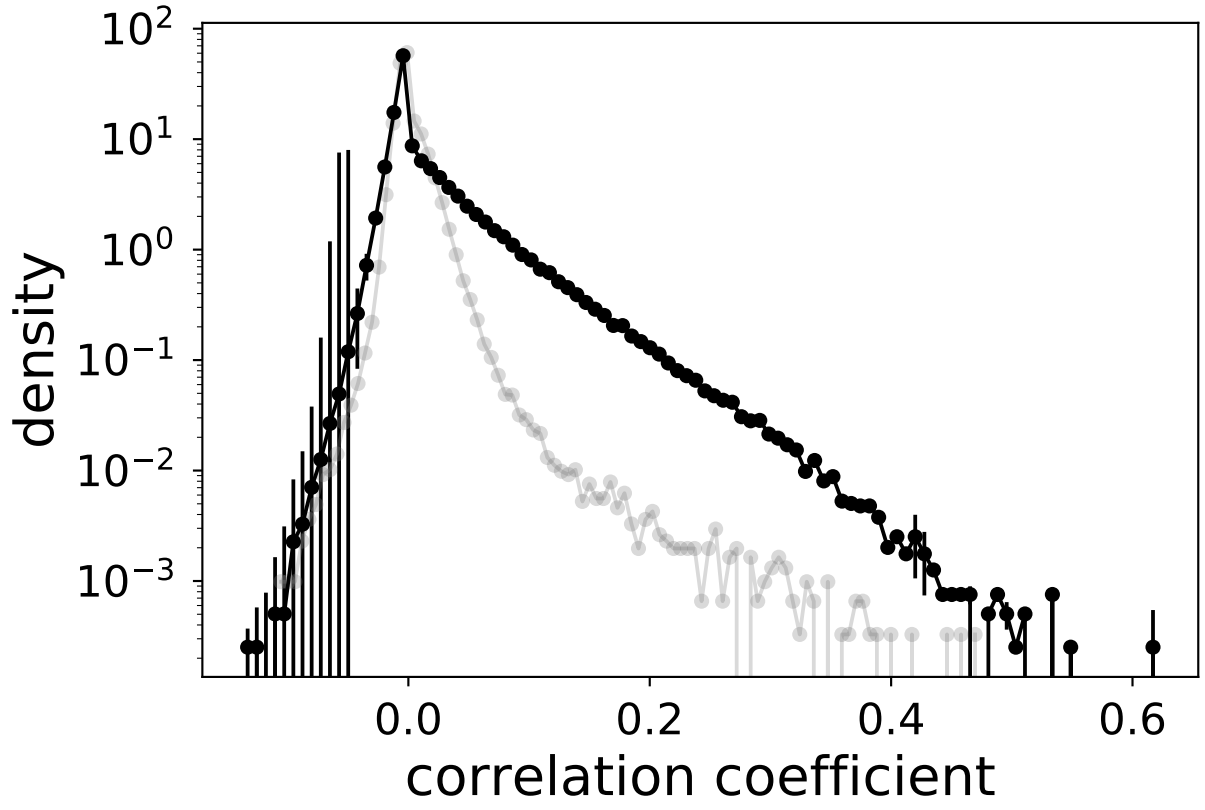

FIG. 5. **(A)** Distribution of pairwise correlation coefficients. **(B)** Rate of firing vs rank of neuron. For **(A)** and **(B)**, error bars are standard deviations over randomly selected contiguous quarters of the simulation.

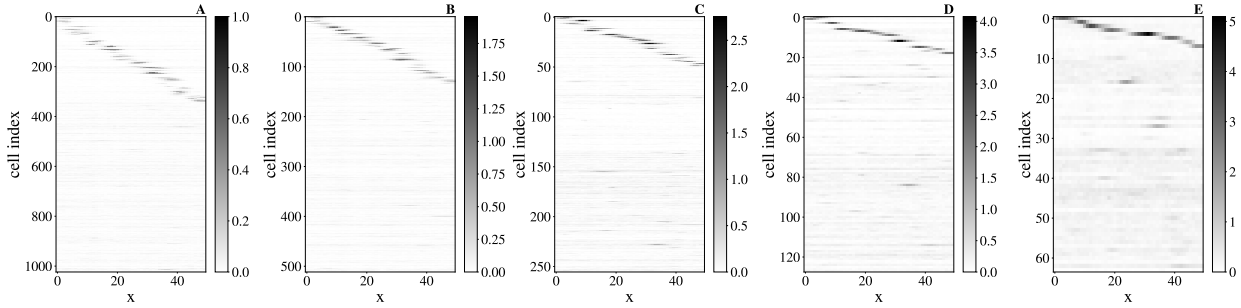

FIG. 6. Average activity at spatial location x for each neuron. Rows are sorted by value of spatial information, greatest to least.

In FIG. 5B we show neuron firing rate vs neuron rank. We note that FIG. 5B is qualitatively similar to the analogous result produced in Ref. [1]. Features reconstructed from Ref. [1] include a maximum firing rate of less than 0.2 and a slight elbow in the otherwise near-straight rate vs rank curve.

Code and documentation regarding this work can be found at *https://github.com/mcmorre*.

B. Parameter sweeps

In order to further investigate the behavior of our model, we perform sweeps over the parameters detailed in TABLE II.

The default parameters, listed in TABLE I result in critical behavior. One by one, we sweep over each parameter while holding all others at default value.

1. Varying cell stimuli

We will now consider the effects of the types of stimuli driving cells on our spin simulation. We complete the following 3 simulations:

1. $N/2$ cells coupled only to place fields, $N/2$ cells coupled to only latent fields
2. N cells coupled only to latent fields
3. $N/2$ cells coupled only to latent fields, $N/2$ cells coupled to both latent fields and place fields.
4. N cells coupled only to place fields.

We refer to these simulations as both, none, place, and no latent, respectively. Note that for the no latent simulation, we increased ε from default value $\varepsilon = -16./6.$ to $\varepsilon = -16./12.$ to account for the omission of latent fields. We then examine neuron activity as we iterate the coarse-graining procedure. In FIG 7 through FIG 15 we show that including place fields in our simulations with latent fields does not significantly alter free energy scaling (FIG 11), correlation time scaling (FIG 13), or approach to a non-gaussian fixed point (FIG 15). However, including place fields in simulations with latent fields creates slight deviation from power law scaling in variance at large cluster size (FIG 10) and is damaging to eigenvalue collapse (FIG 9.

Omitting latent fields from simulations has a disastrous effect on scaling. In FIG 7 through FIG 15 we show that including place fields only in our simulations significantly alters free energy scaling (FIG 11), correlation time scaling (FIG 13), approach to a non-gaussian fixed point (FIG 15), power law scaling at large cluster size (FIG 10), and eigenvalue collapse (FIG 9.

By examining FIG 7 through FIG 15, we can conclude that the presence of scaling behavior does not depend on the presence of place fields, but does depend on whether all cells also are

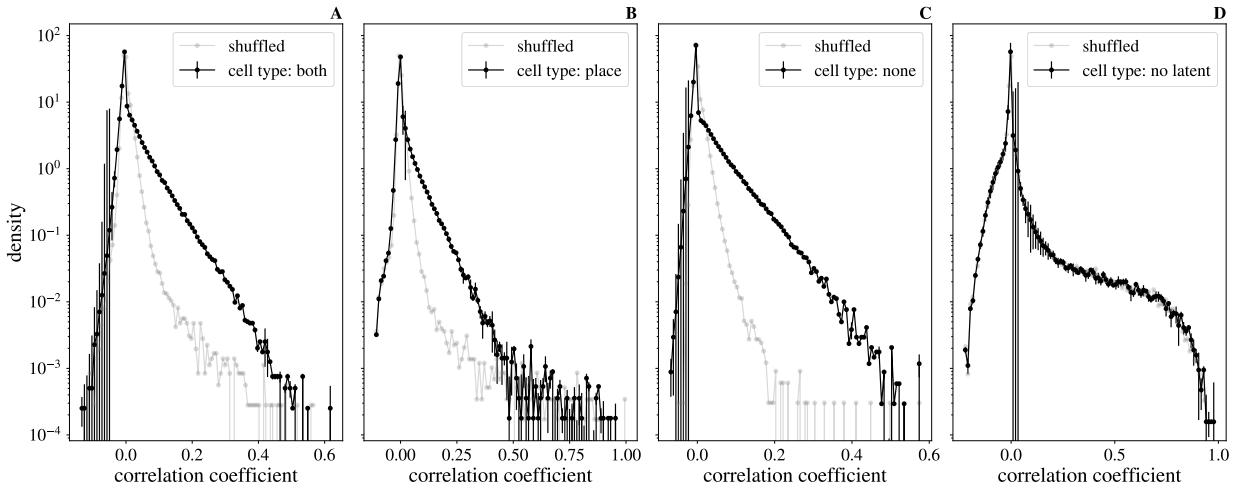

FIG. 7. Distributions of pairwise correlation coefficients. Default simulation parameters with labeled simulation type. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

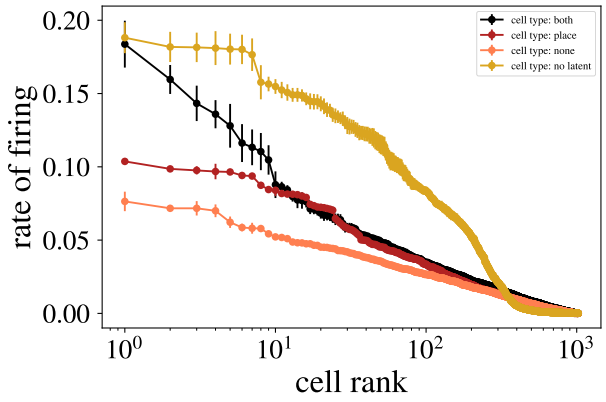

FIG. 8. Rate of firing vs rescaled rank. Default simulation parameters with labeled simulation type. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

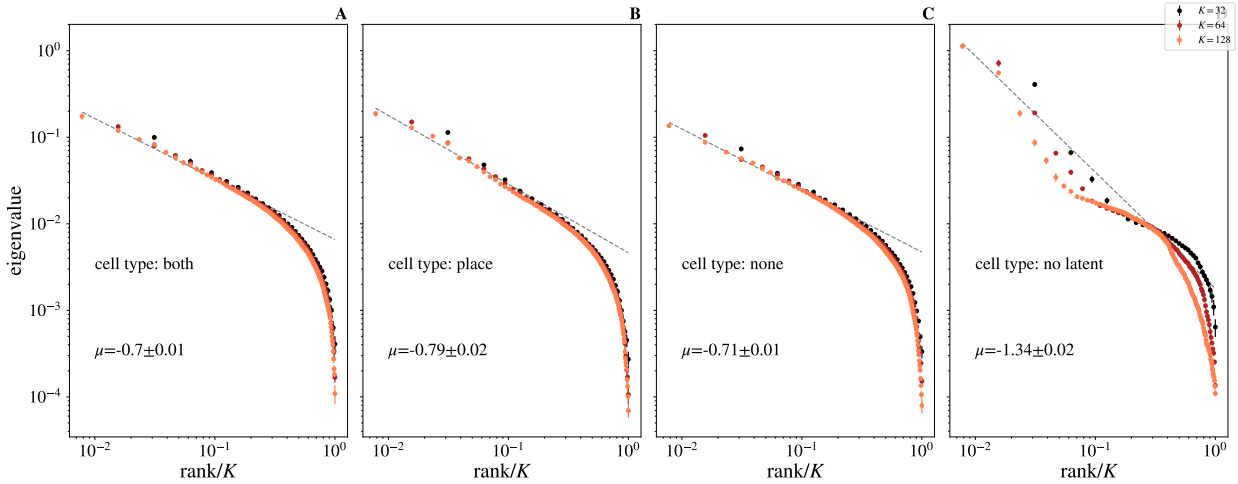

FIG. 9. Average eigenvalue spectrum of Eq 40 for cluster sizes $K = 32, 64, 128$. Default simulation parameters with labeled simulation type. No significant difference in quality of scaling is observed for the both and place simultions, while the place simulation has a lower quality eigenvalue collapse. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

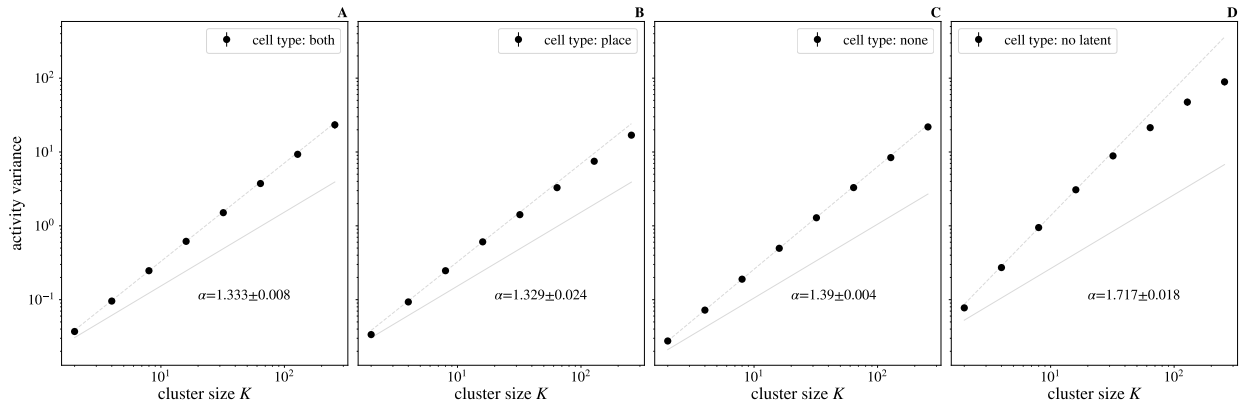

FIG. 10. Activity variance (Eq 33) over coarse-grained variables at each coarse-graining iteration. No significant difference in quality of scaling is observed for the both and place simultions, while the place simulation has a lower quality variance scaling. Default simulation parameters with labeled simulation type. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

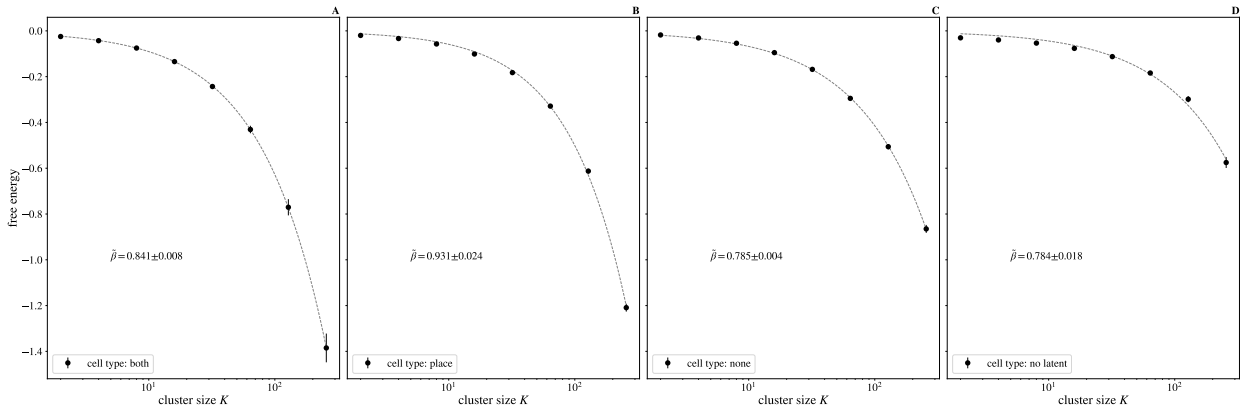

FIG. 11. Average free energy (Eq 34) at each coarse-graining iteration. Default simulation parameters with labeled simulation type. No significant differences in quality of scaling between simulations is observed. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

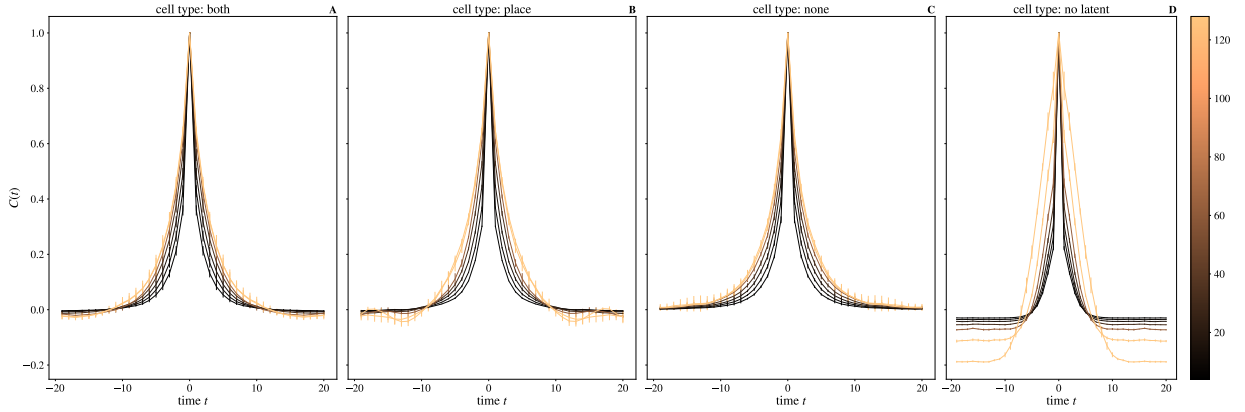

FIG. 12. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$ as a function of time, cluster size indicated by colorbar. Default simulation parameters with labeled simulation type. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

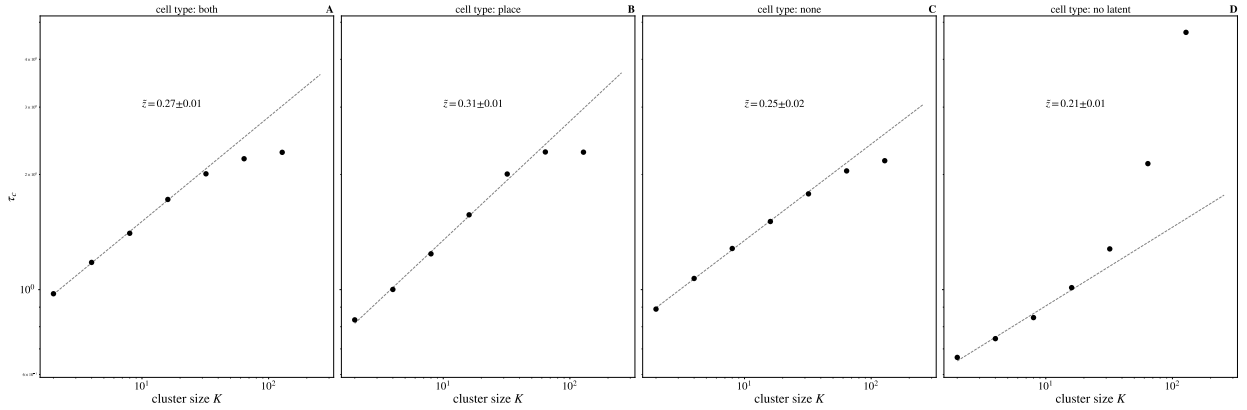

FIG. 13. Time constants τ_c extracted from each curve in in FIG 12, and observe behavior obeying $\tau_c \propto K^{\tilde{z}}$ for roughly 1 decade. Default simulation parameters with labeled simulation type. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

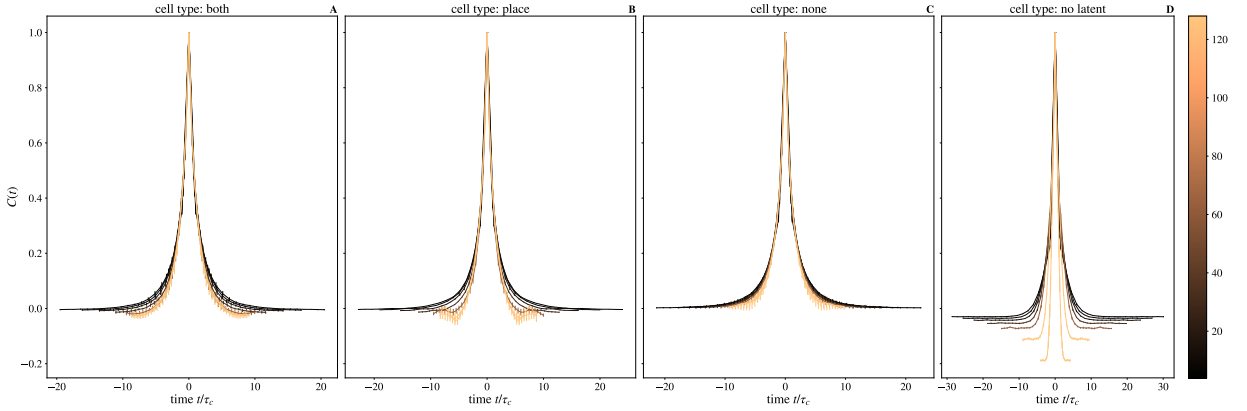

FIG. 14. Average autocorrelation function for cluster sizes $K = 2,4,...,256$, where time is rescaled by the appropriate τ_c for that coarse-graining iteration. Default simulation parameters with labeled simulation type. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

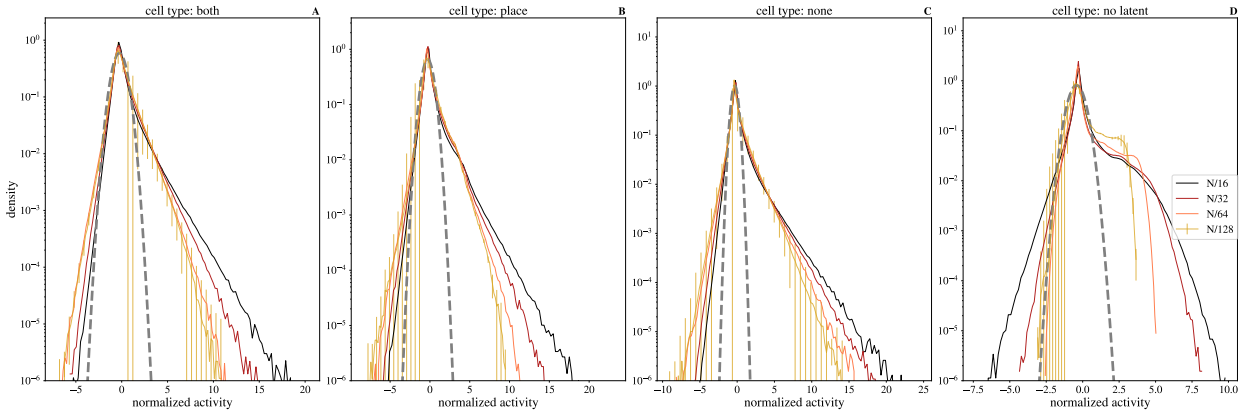

FIG. 15. Distribution of coarse-grained variables for $k = N/16, N/32, N/64, N/128$ modes retained. No significant differences in approach to a non-gaussian fixed point are observed between simulation types. Default simulation parameters with labeled simulation type. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

coupled to latent stimuli. In fact, the presence of place fields is deleterious to scaling and does not yield scaling behavior without the inclusion of latent fields. This is expected, since we propose that the only two requirements to construct criticality and therefore scaling behavior are 1) large system size and 2) the system is driven by latent dynamic stimuli.

2. Varying the latent fields multiplier ϕ

We now vary the latent fields multiplier ϕ. We perform simulations with the default parameters sweeping over values of ϕ ranging from $\phi = 0.8$ to $\phi = 1.5$. In FIG 26, we note that our simulations include a regime which quantitatively matches experimental results [1].

In FIG 16 through FIG 23 we show that the presence of place cells stays stable over coarse graining in the range $\phi \in [0.9, 1.2]$. We find that varying ϕ does not significantly effect eigenvalue scaling results (FIG 27, free energy scaling (FIG 29), correlation time scaling (FIG 31), or approach to a non-gaussian fixed point (FIG 33).

However, setting $\phi < 1.0$ creates slight deviation from power law scaling at large cluster size (FIG 28). In summary, FIG 16 through FIG 33 show that latent fields which are too weak or too strong destroy the presence of place cells over coarse-graining iterations, and latent fields which are too weak are deleterious to scaling behavior.

3. Varying the number of latent fields n_f

We will now consider the effects of varying the number of latent fields n_f in our spin simulation. We perform simulations with the default parameters sweeping over values of n_f ranging from $n_f = 1$ to $n_f = 20$. In FIG 36, we note that our simulations include a regime which quantitatively matches experimental results [1].

We find that varying n_f significantly damages eigenvalue scaling (FIG 37 and variance scaling (FIG 38) for $n_f < 5$. However, free energy scaling (FIG 39) and correlation time scaling (FIG 41) are not significantly affected by variation in n_f. FIG 34 through FIG 43 suggest that 5 or more latent fields are required to observe scaling behavior.

It is possible there is an upper limit of n_f for a simulation to display critical behavior. As n_f increases, distributions of coarse-grained activity become increasingly short-tailed (FIG 43). In addition, FIG 36A shows variance scaling exponent α approaching 1.2. It is possible for some

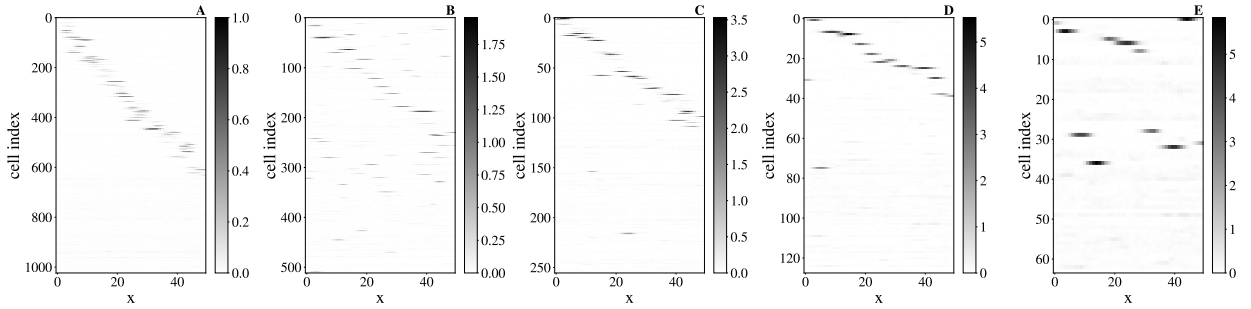

FIG. 16. Average activity at spatial location x for each neuron. Default simulation parameters with $\phi = 0.8$. Coarse-graining steps $0, 1, 2, 3, 4$ displayed by **(A)**, **(B)**, **(C)**, **(D)**, **(E)**, **(F)**, respectively.

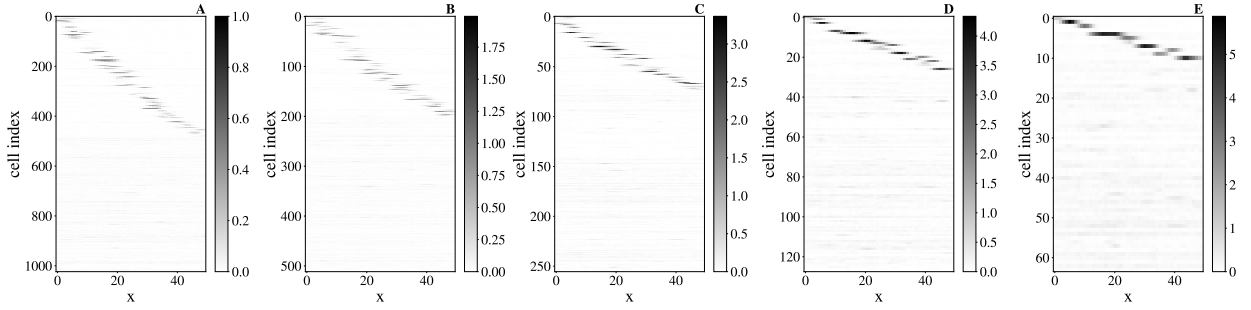

FIG. 17. Average activity at spatial location x for each neuron. Default simulation parameters with $\phi = 0.9$. Coarse-graining steps $0, 1, 2, 3, 4$ displayed by **(A)**, **(B)**, **(C)**, **(D)**, **(E)**, **(F)**, respectively.

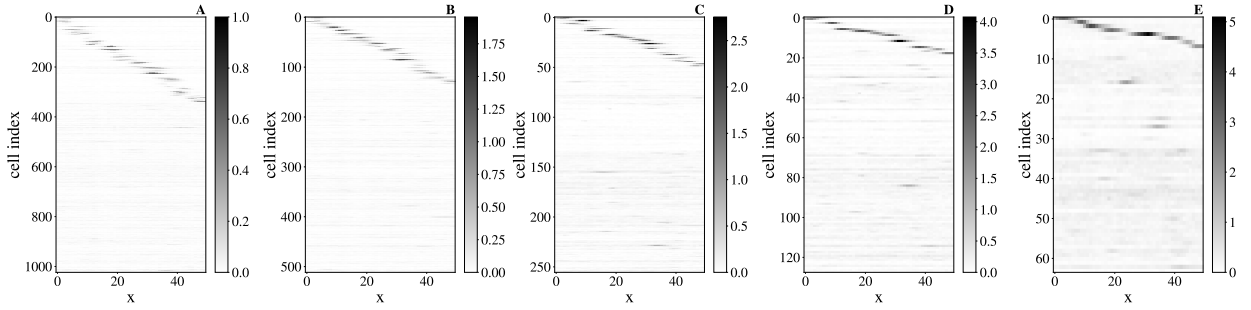

FIG. 18. Average activity at spatial location x for each neuron. Default simulation parameters. Coarse-graining steps $0, 1, 2, 3, 4$ displayed by **(A)**, **(B)**, **(C)**, **(D)**, **(E)**, **(F)**, respectively.

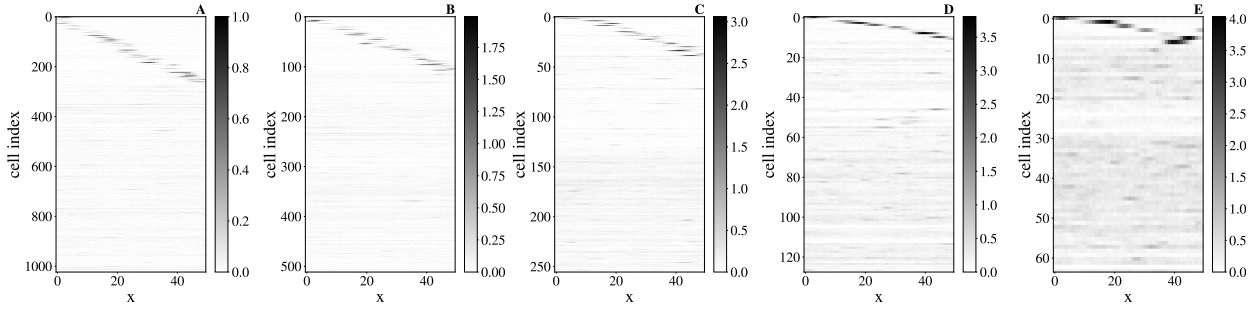

FIG. 19. Average activity at spatial location x for each neuron. Default simulation parameters with $\phi = 1.1$. Coarse-graining steps $0, 1, 2, 3, 4$ displayed by **(A)**, **(B)**, **(C)**, **(D)**, **(E)**, **(F)**, respectively.

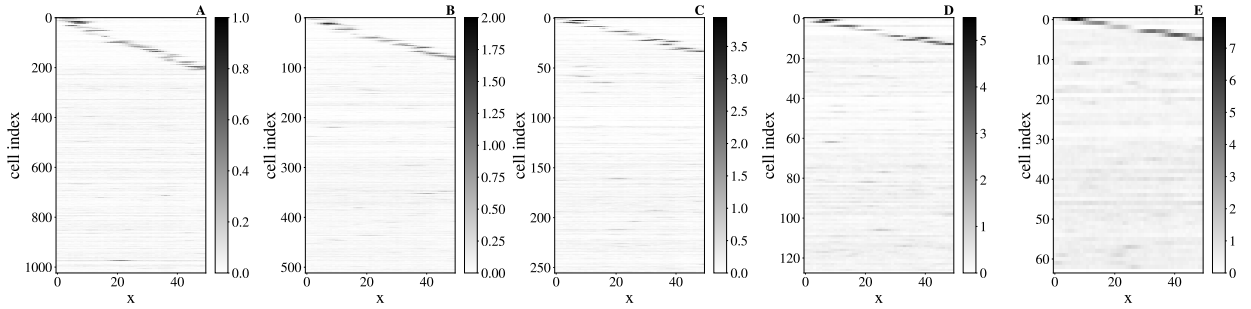

FIG. 20. Average activity at spatial location x for each neuron. Default simulation parameters with $\phi = 1.2$. Coarse-graining steps $0, 1, 2, 3, 4$ displayed by **(A)**, **(B)**, **(C)**, **(D)**, **(E)**, **(F)**, respectively.

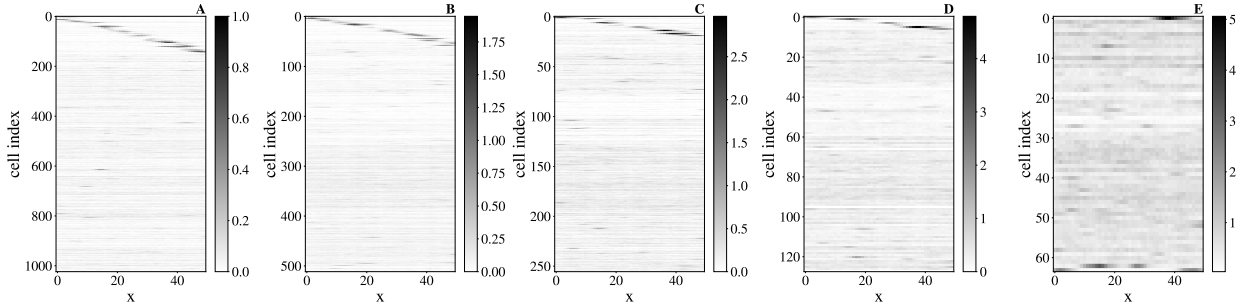

FIG. 21. Average activity at spatial location x for each neuron. Default simulation parameters with $\phi = 1.3$. Coarse-graining steps $0, 1, 2, 3, 4$ displayed by **(A)**, **(B)**, **(C)**, **(D)**, **(E)**, **(F)**, respectively.

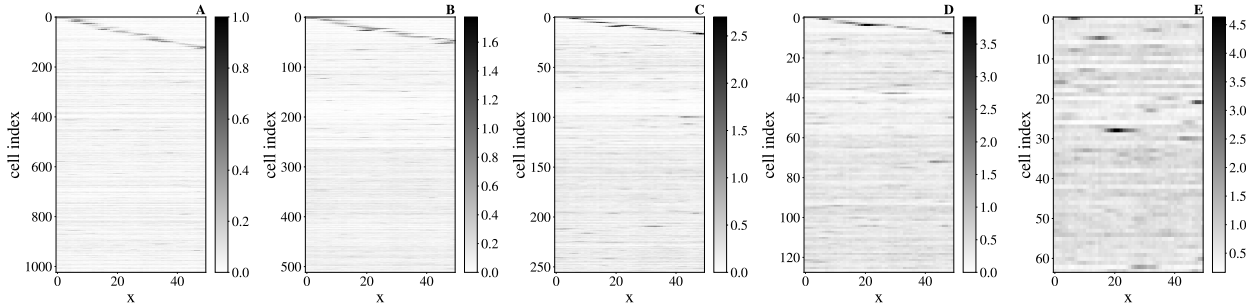

FIG. 22. Average activity at spatial location x for each neuron. Default simulation parameters with $\phi = 1.4$. Coarse-graining steps $0, 1, 2, 3, 4$ displayed by **(A)**, **(B)**, **(C)**, **(D)**, **(E)**, **(F)**, respectively.

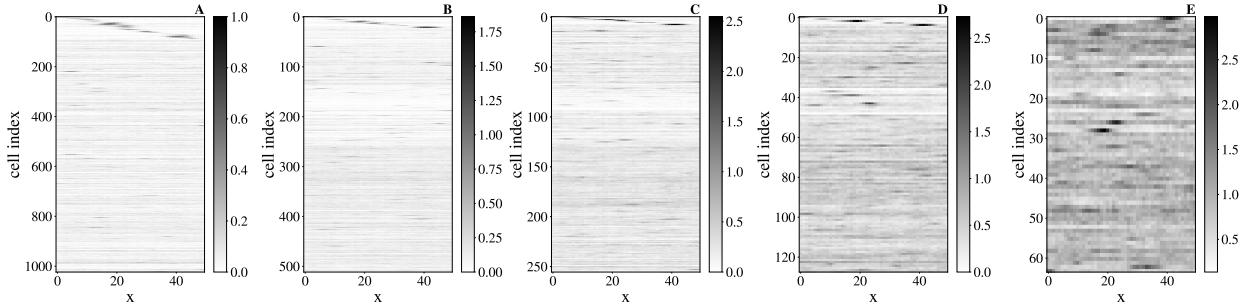

FIG. 23. Average activity at spatial location x for each neuron. Default simulation parameters with $\phi = 1.5$. Coarse-graining steps $0, 1, 2, 3, 4$ displayed by **(A)**, **(B)**, **(C)**, **(D)**, **(E)**, **(F)**, respectively.

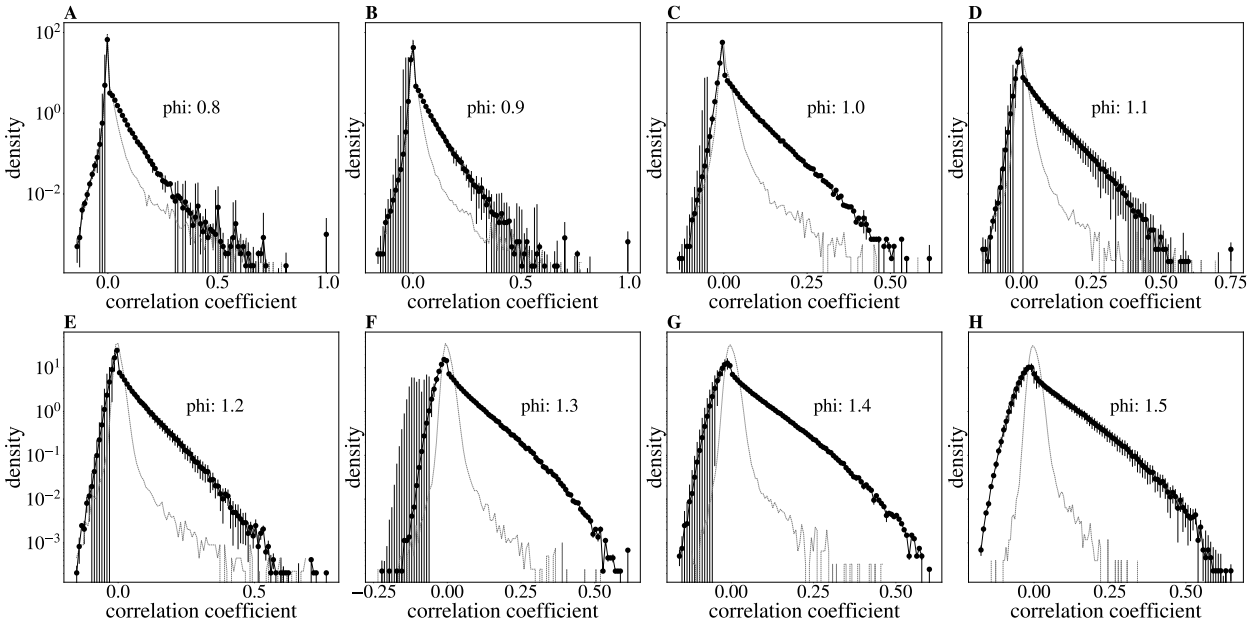

FIG. 24. Distributions of pairwise correlation coefficients. Default simulation parameters with value of ϕ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

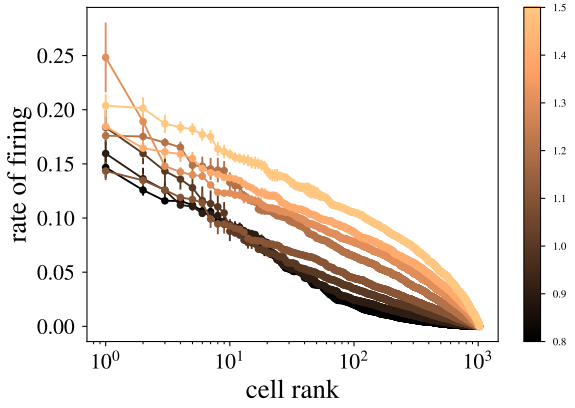

FIG. 25. Rate of firing vs rescaled rank. Colorbar indicates value of ϕ

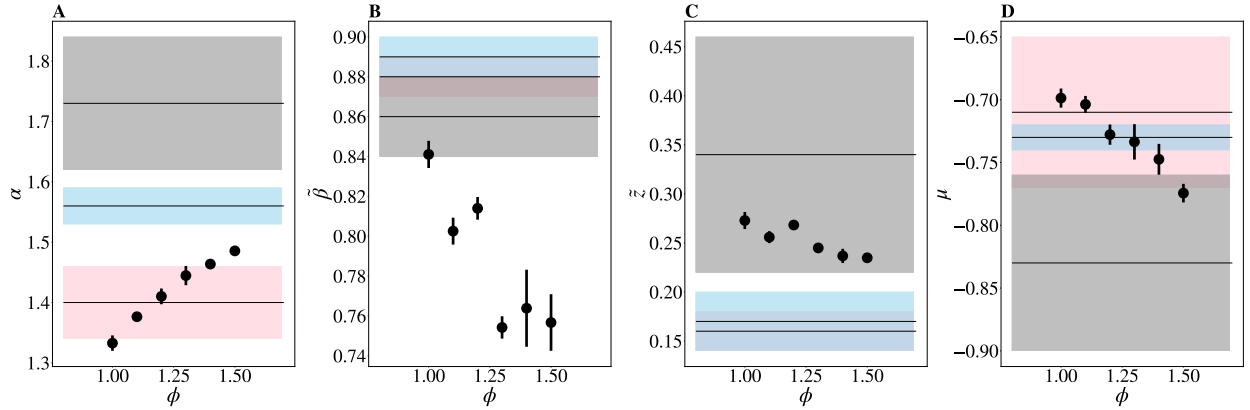

FIG. 26. Each critical exponent, $\alpha, \tilde{\beta}, \tilde{z}, \mu$ vs latent field multiplier ϕ. Results from [1] marked and shaded in gray, pink, and blue. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

$n_f > 20$, $\alpha \approx 1$ and the simulation will become non-critical. More simulations need to be performed to confirm this hypothesis.

4. Varying the latent field time constant τ

We will now consider the effects of varying τ in our simulation. We perform simulations with the default parameters sweeping over values of τ. In FIG 46, we note that our simulations include a regime which quantitatively matches experimental results [1].

We find that varying τ causes deviations in the exponent $\tilde{z}$, with larger τ corresponding to larger $\tilde{z}$ and smaller τ corresponding to small $\tilde{z}$ (FIG 46). However, free energy scaling (FIG 49), variance scaling (FIG 48), eigenvalue scaling (FIG 47, and approach to a non-gaussian fixed point (FIG 43) are not significantly affected by variation in τ_f. FIG 44 through FIG 53 suggest that dynamic scaling results are not significantly improved by increasing correlation time of the latent fields.

5. Varying the probability of coupling to a latent field q

We will now consider the effects of varying q in our spin simulation. We perform simulations with the default parameters sweeping over values of q. In FIG 66, we note that our simulations include a regime which quantitatively matches experimental results [1].

We find that varying q causes slight deviations in variance scaling for $q < 0.5$. Free energy scaling (FIG 59), variance scaling (FIG 58), eigenvalue scaling (FIG 57), approach to a non-gaussian fixed point (FIG 63), and correlation time scaling (FIG 41) are not significantly affected by variation in q. We conclude that varying the probability of coupling to a latent field does not have a significant impact on scaling for $q \geq 0.5$, but is deleterious to scaling for $q < 0.5$.

6. Varying the penalty term ε

We perform simulations with the default parameters sweeping over values of ε. We find that varying ε significantly damages eigenvalue scaling for $\varepsilon > -1.92$ (FIG 67). In FIG 66, we note that our simulations include a regime which quantitatively matches experimental results [1].

We observe that coarse-grained distributions of activity from simulations with $\varepsilon > -1.5$ ap-

proach but to dot reach a gaussian fixed point (FIG 73). However, free energy scaling (FIG 69), variance scaling (FIG 68), and correlation time scaling (FIG 41) are not significantly affected by variation in ε. We conclude that highly active simulations do not display perfect eigenvalue spectra collapse.

7. *Varying the multiplier* η

We perform simulations with the default parameters sweeping over values of η. In FIG 76, we note that our simulations include a regime which quantitatively matches experimental results [1].

Free energy scaling (FIG 79), variance scaling (FIG 78), eigenvalue scaling (FIG 77), approach to a non-gaussian fixed point (FIG 83), and dynamic scaling are not significantly affected by variation in η. Thus, adjusting η appears to have no effect on scaling results.

---

[1] L. Meshulam, J. L. Gauthier, C. D. Brody, D. W. Tank, and W. Bialek, Coarse–graining and hints of scaling in a population of 1000+ neurons, arXiv:1812.11904. (2018).

[2] K. Huang, *Statistical Mechanics*, 2nd ed. (John Wiley & Sons, 1987).

[3] L. P. Kadanoff, Scaling laws for ising models near T_c, Physics Physique Fizika **2**, 263 (1966).

[4] K. G. Wilson and J. Kogut, The renormalization group and the expansion, Physics Reports **12**, 75 (1974).

[5] T. Mora and W. Bialek, Are biological systems poised at criticality?, J Stat Phys **144**, 268–302 (2011).

[6] D. J. Schwab, I. Nemenman, and P. Mehta, Zipf's law and criticality in multivariate data without fine-tuning, Phys. Rev. Lett. **113**, 068102 (2014).

[7] J. O'Keefe and J. Dostrovsky, The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat, Brain Research **34**, 171 (1971).

[8] L. Meshulam, J. L. Gauthier, C. D. Brody, D. W. Tank, and W. Bialek, Collective behavior of place and non-place neurons in the hippocampal network, Neuron **96**, 1178 (2017).

[9] E. Schneidman, M. Berry II, R. Segev, and W. Bialek, Weak pairwise correlations imply strongly correlated network states in a neural population, Nature **440**, 1007 (2006).

[10] R. M. Gray, Toeplitz and circulant matrices: A review, Foundations and Trends® in Communications and Information Theory **2**, 155 (2006).

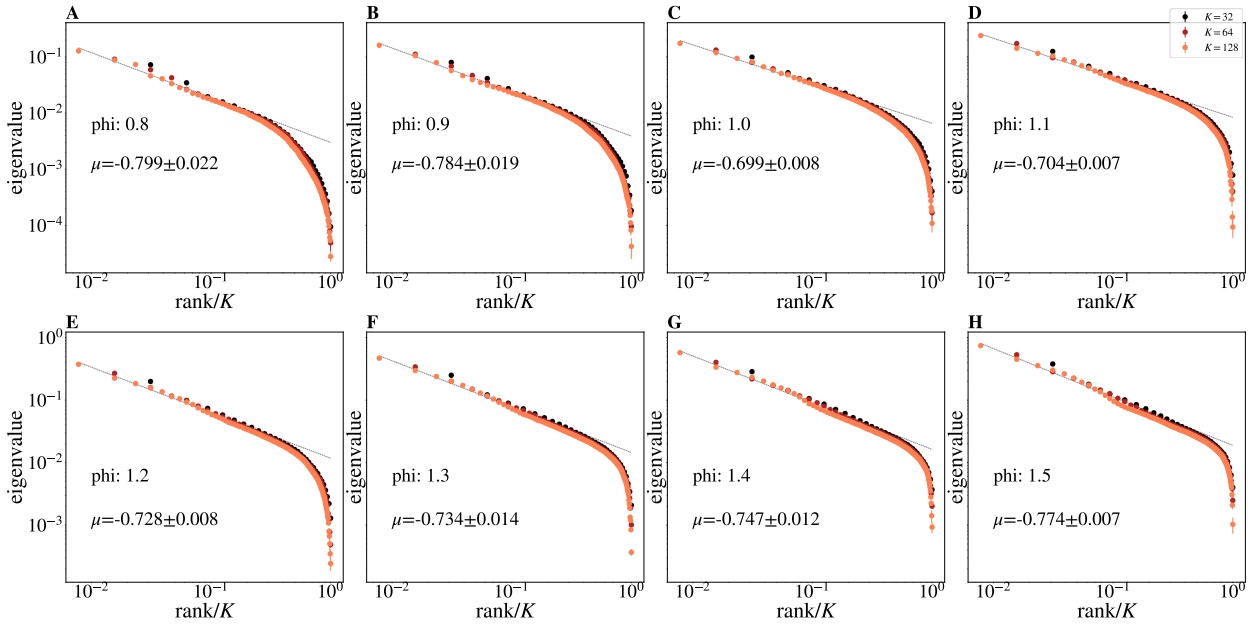

FIG. 27. Average eigenvalue spectrum of Eq 40 for cluster sizes $K = 32, 64, 128$. Default simulation parameters with labeled values of ϕ. Note that quality of eigenvalue collapse and scaling is unaffected by value of ϕ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

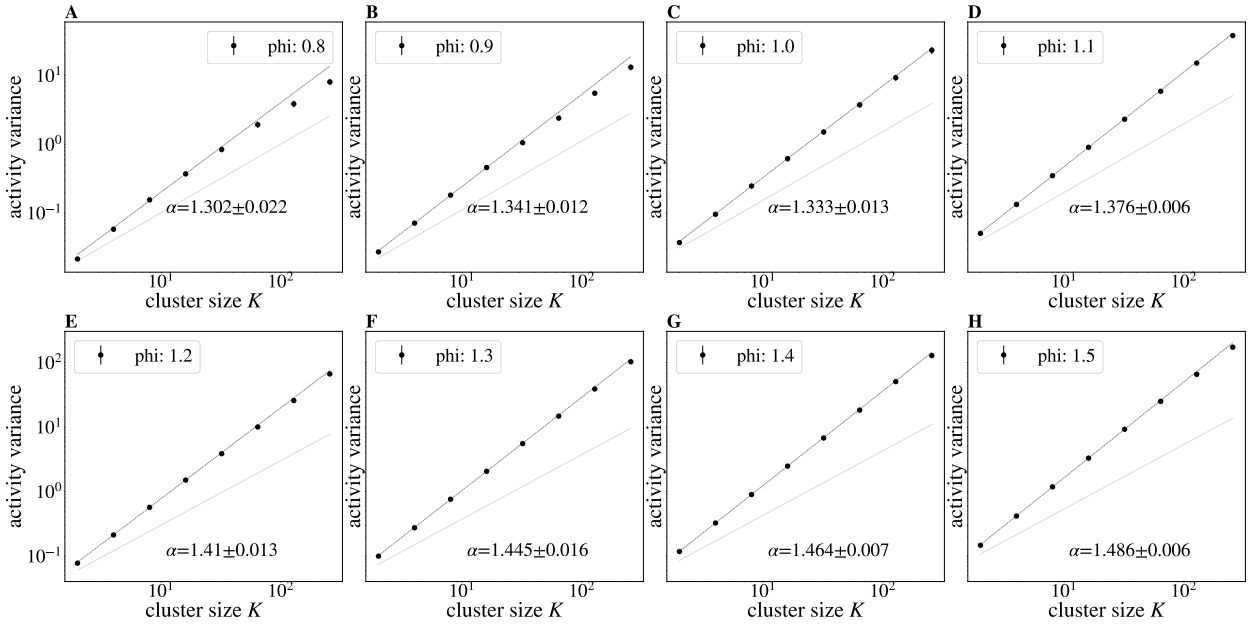

FIG. 28. Activity variance (Eq 33) over coarse-grained variables at each coarse-graining iteration. Default simulation parameters with labeled values of ϕ. Note that for $\phi < 1.0$, variance scaling is damaged. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

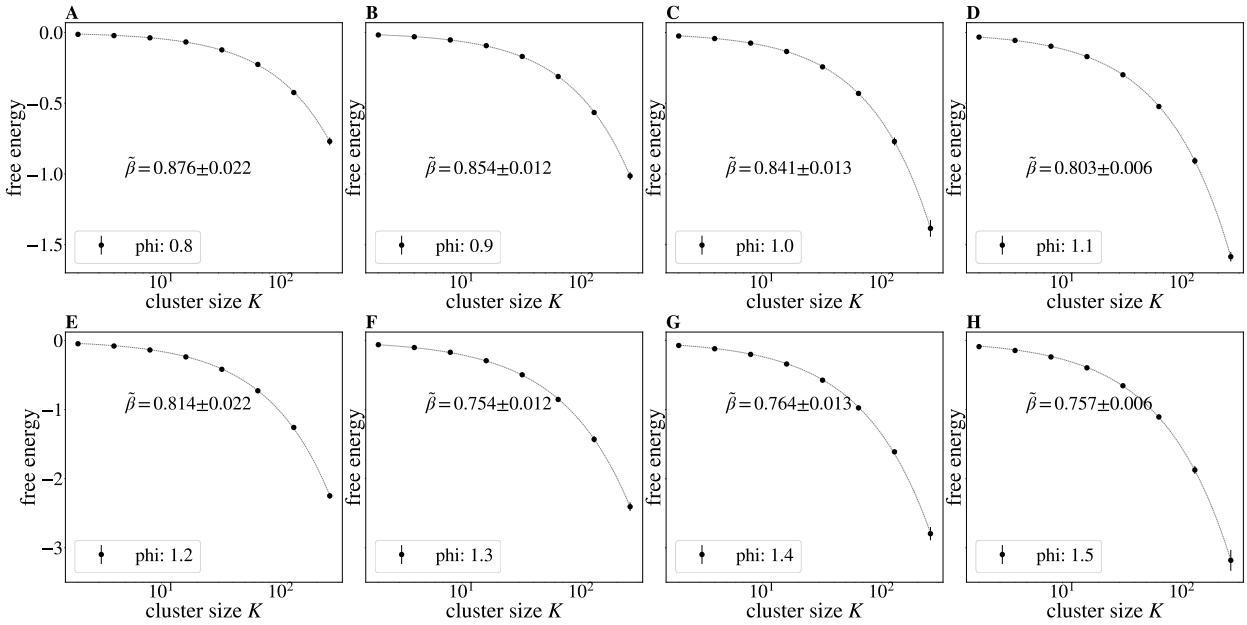

FIG. 29. Average free energy (Eq 34) at each coarse-graining iteration. Default simulation parameters with labeled values of ϕ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

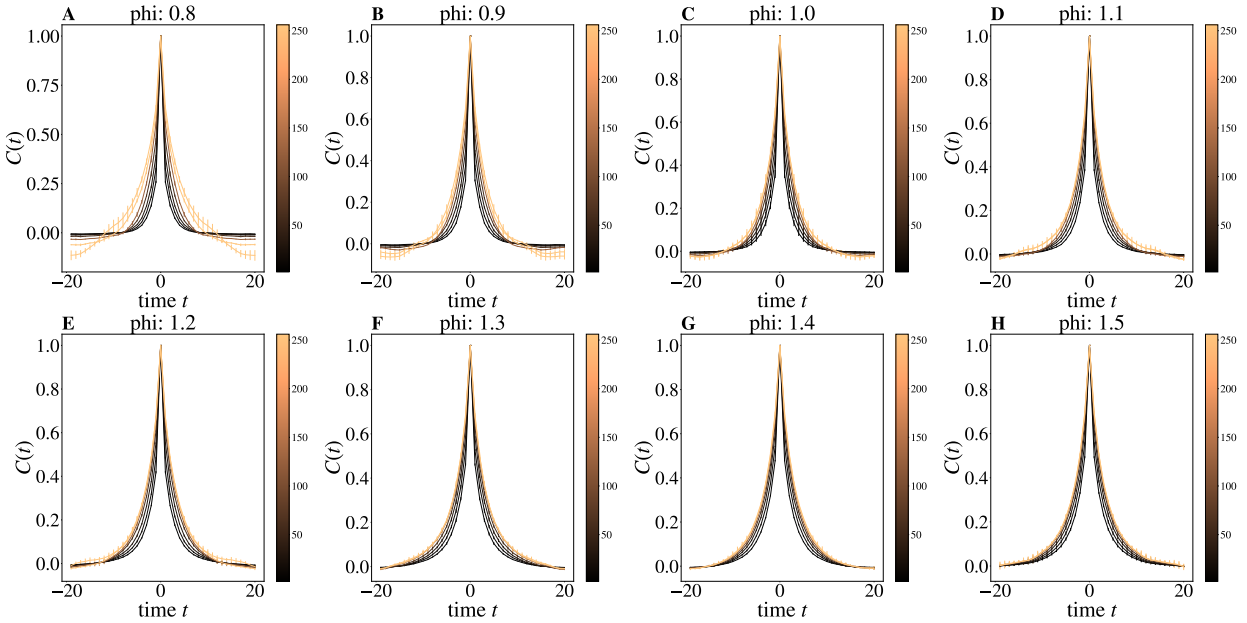

FIG. 30. Average autocorrelation function for cluster sizes $K = 2, 4, \ldots, 256$ as a function of time, cluster size indicated by colorbar. Default simulation parameters with labeled values of ϕ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

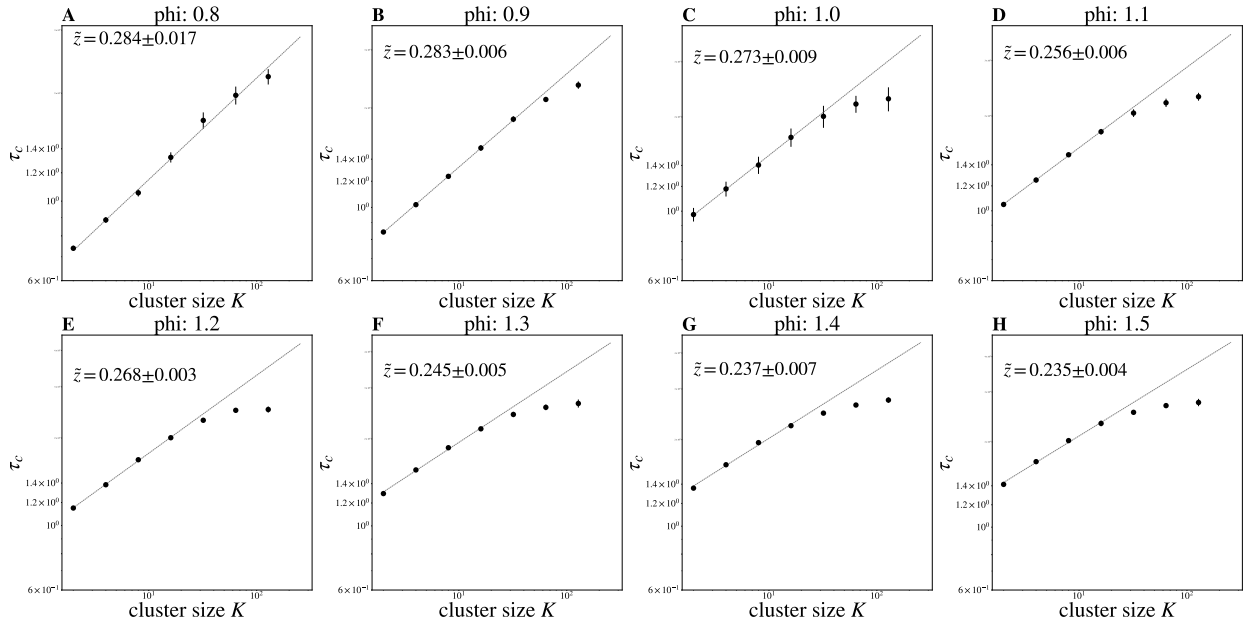

FIG. 31. Time constants τ_c extracted from each curve in in FIG 30, and observe behavior obeying $\tau_c \propto K^{\tilde{z}}$ for roughly 1 decade. Default simulation parameters with labeled values of ϕ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

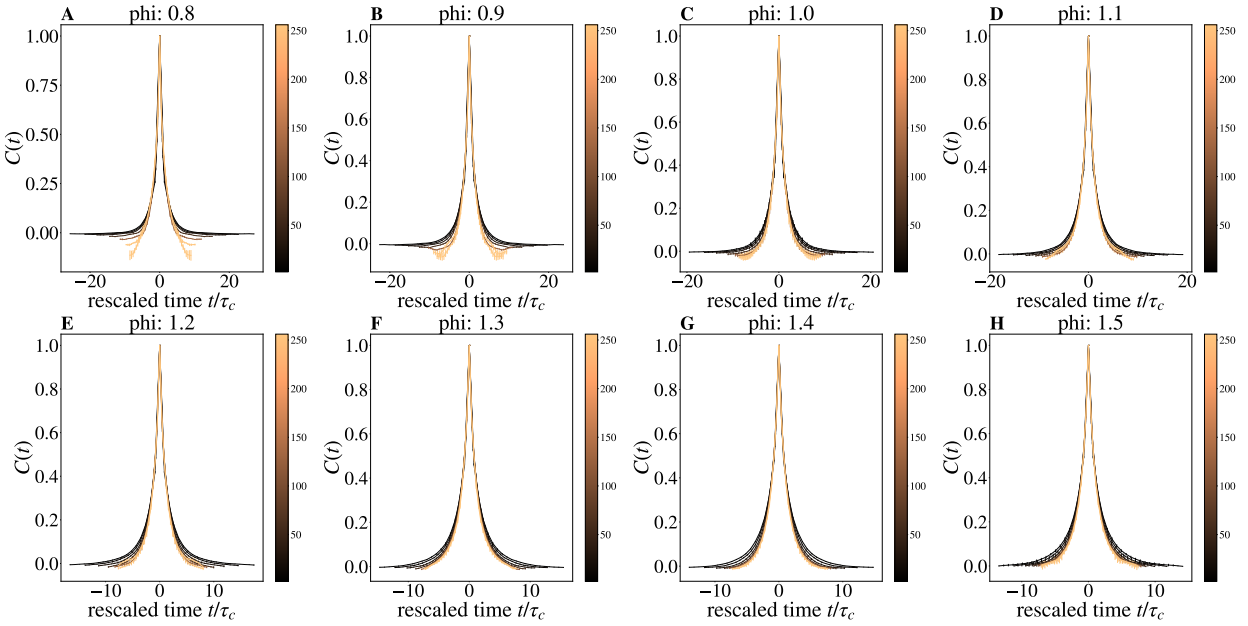

FIG. 32. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$, where time is rescaled by the appropriate τ_c for that coarse-graining iteration. Default simulation parameters with labeled values of ϕ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

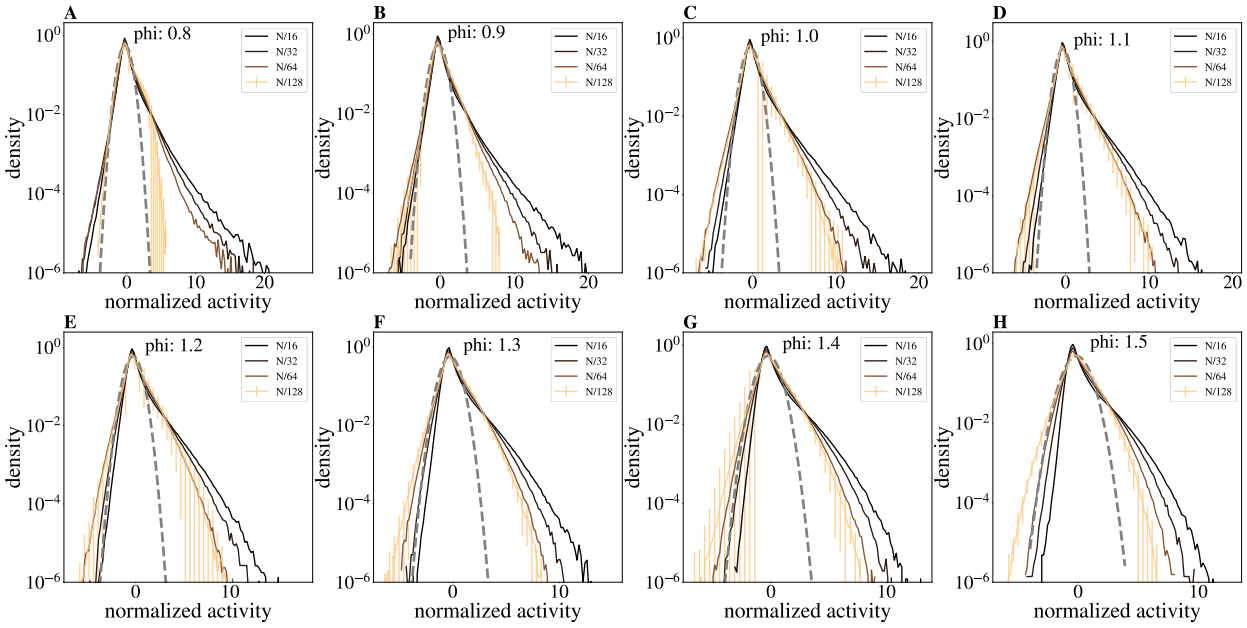

FIG. 33. Distribution of coarse-grained variables for $k = N/16, N/32, N/64, N/128$ modes retained. Note convergence to non-gaussian fixed point regardless of value of ϕ. Default simulation parameters with labeled values of ϕ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

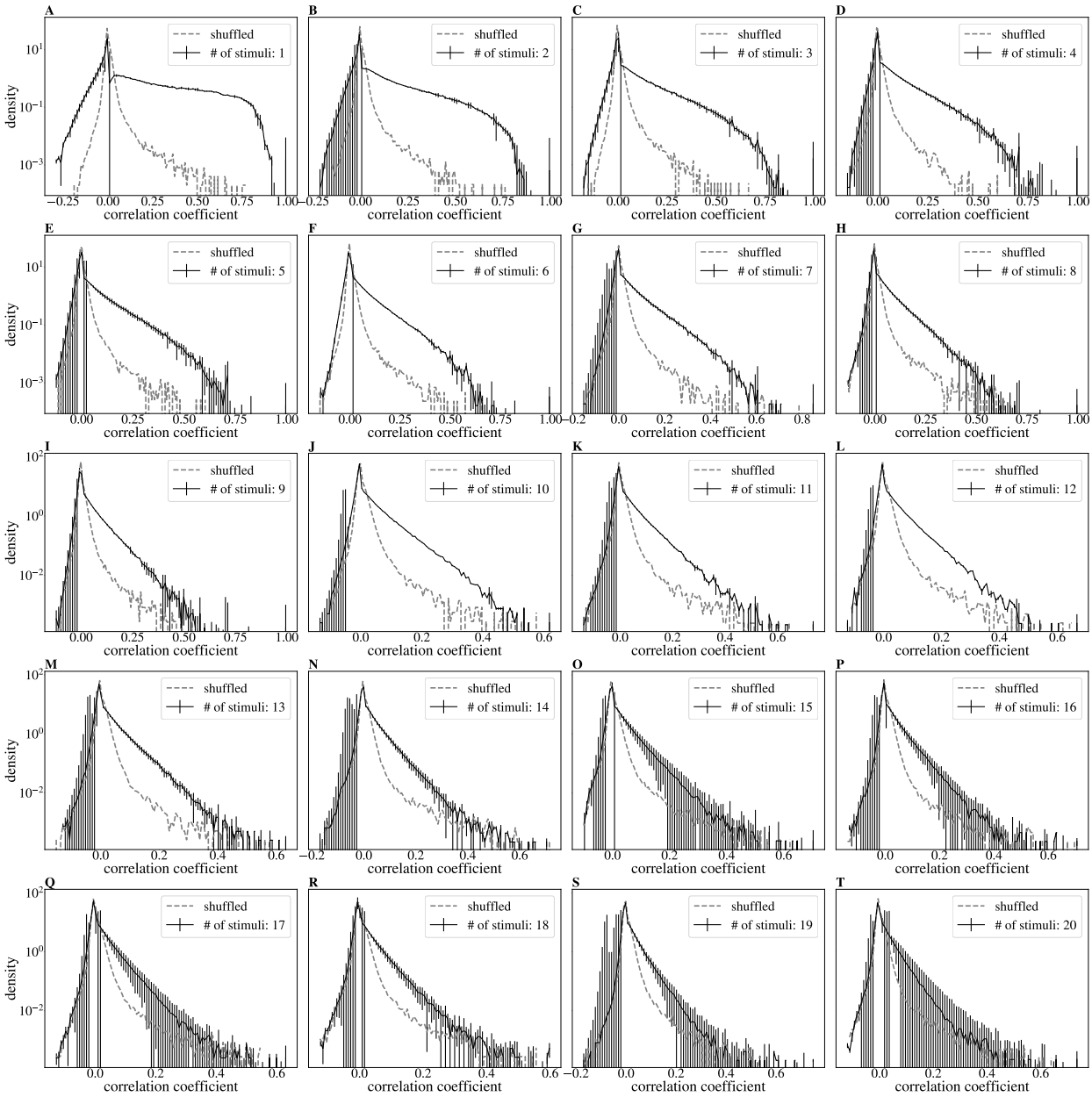

FIG. 34. Distributions of pairwise correlation coefficients. Default simulation parameters with value of n_f. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

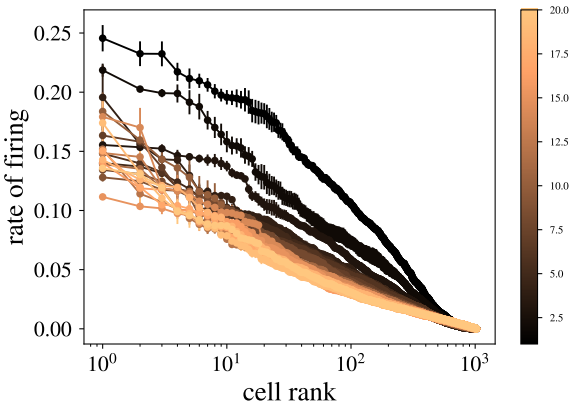

FIG. 35. Rate of firing vs rescaled rank. Colorbar indicates value of n_f. Setting $n_f < 4$ creates a steeper curve.

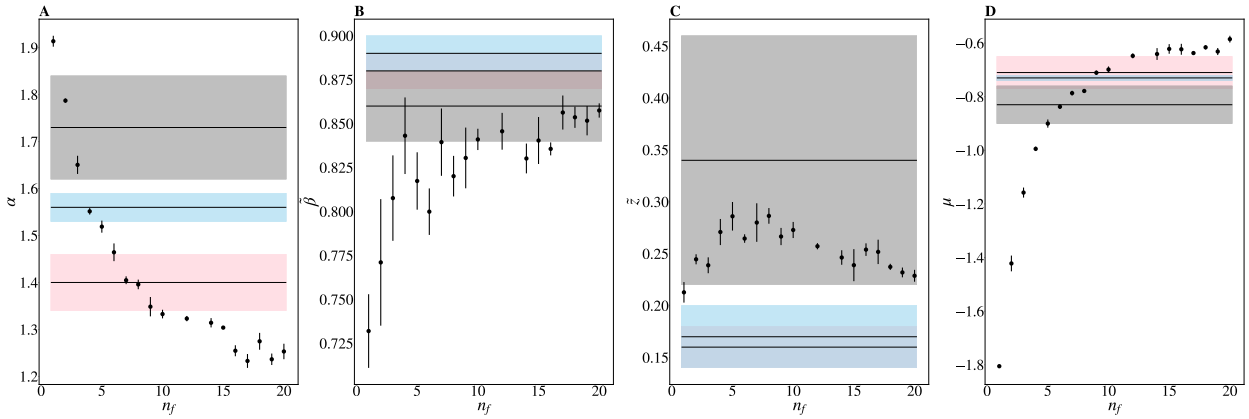

FIG. 36. Each critical exponent, $\alpha, \tilde{\beta}, \tilde{z}, \mu$ vs number of latent fields n_f. Results from experiment [1] marked and shaded in gray, pink, and blue. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

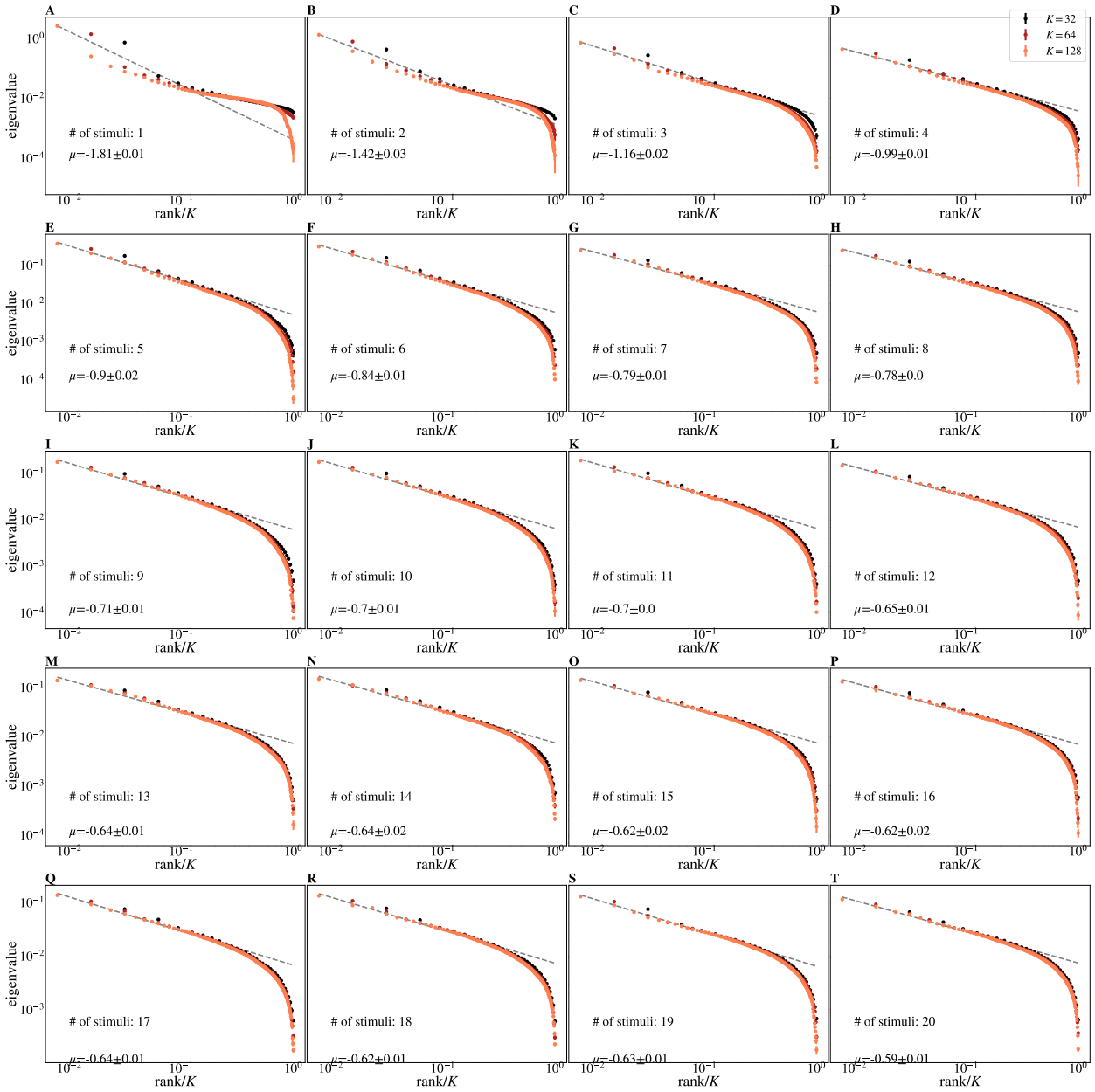

FIG. 37. Average eigenvalue spectrum of Eq 40 for cluster sizes $K = 32, 64, 128$. Default simulation parameters with labeled values of n_f. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

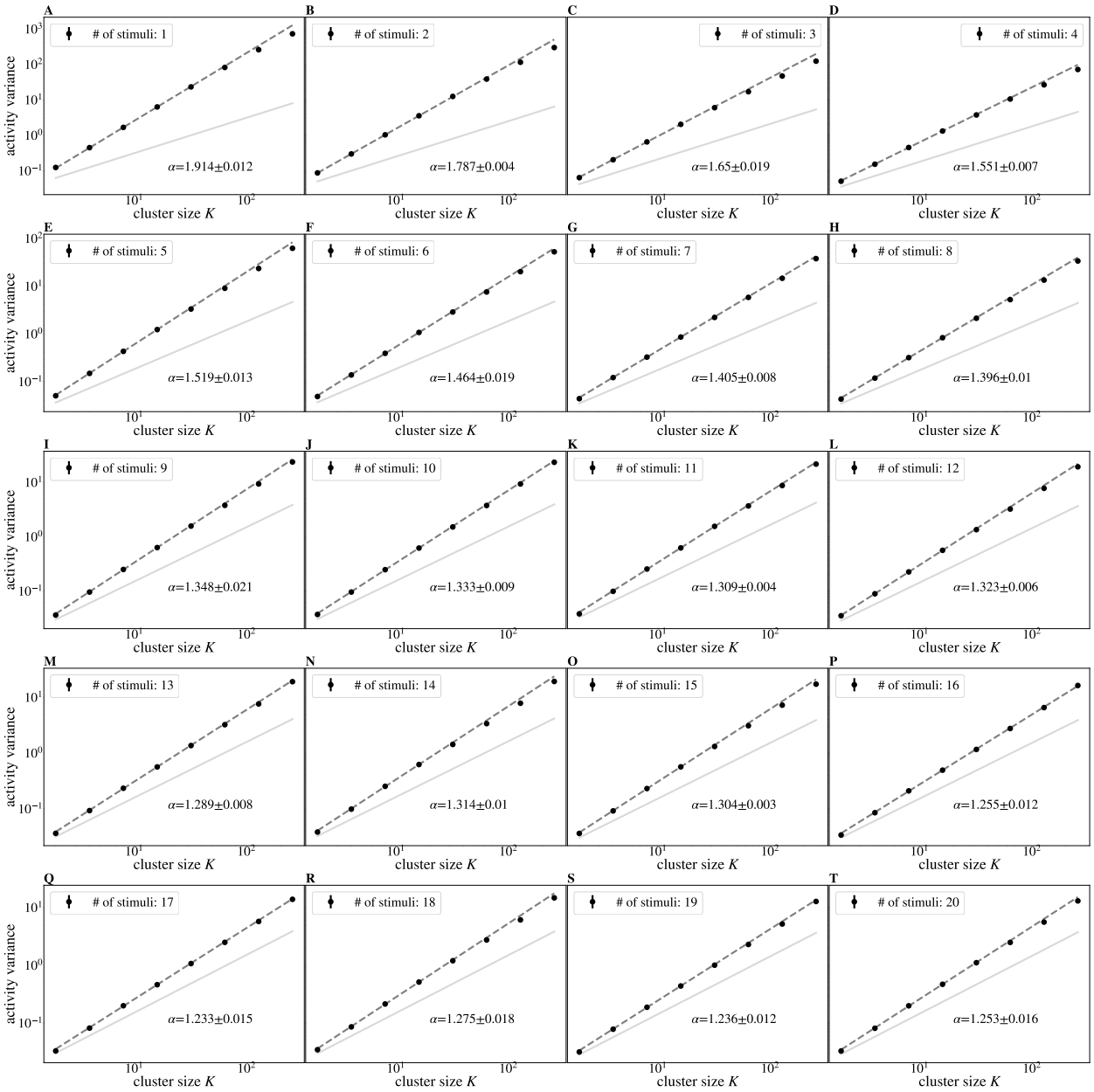

FIG. 38. Activity variance (Eq 33) over coarse-grained variables at each coarse-graining iteration. Default simulation parameters with labeled values of n_f.

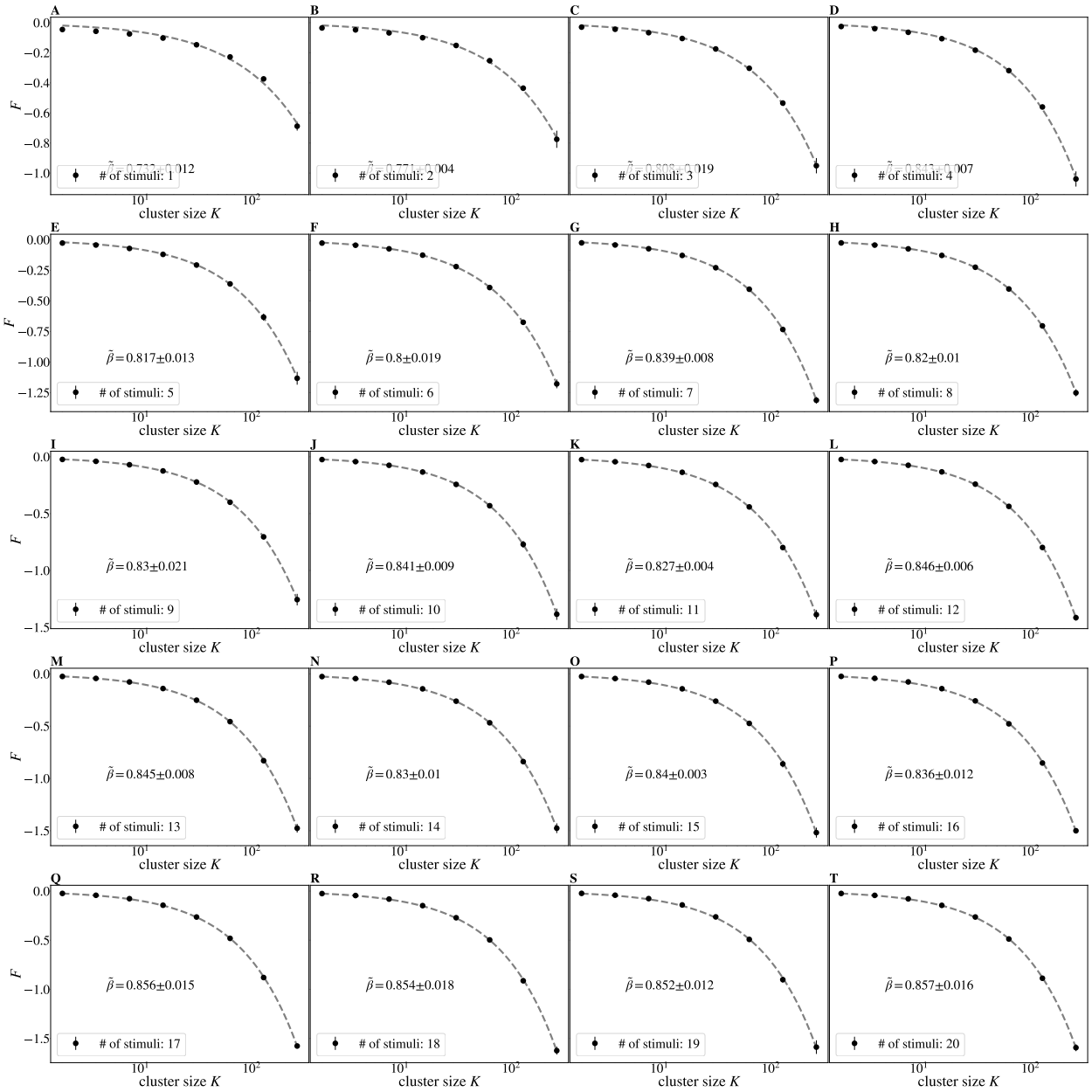

FIG. 39. Average free energy (Eq 34) at each coarse-graining iteration. Default simulation parameters with labeled values of n_f.

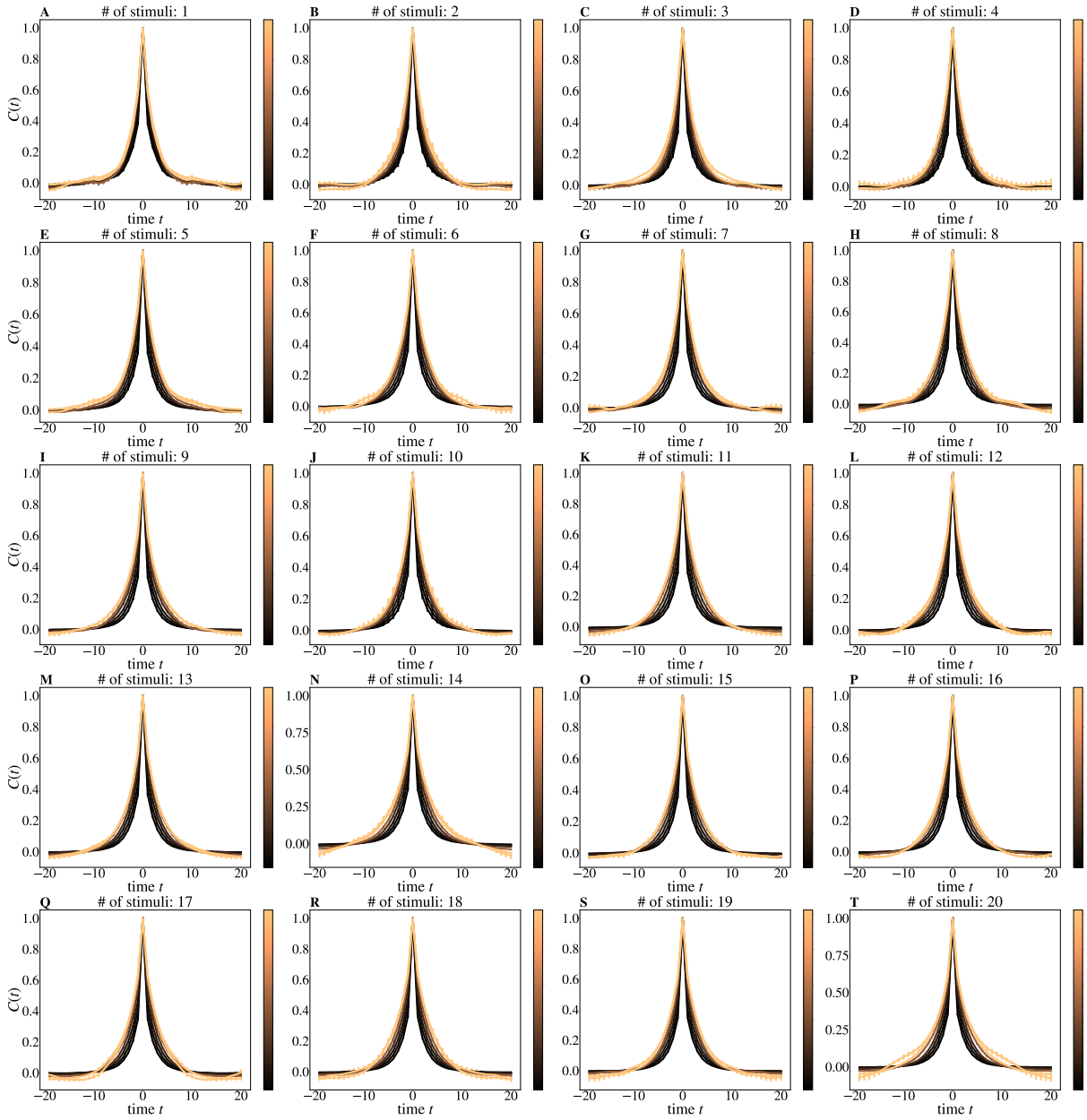

FIG. 40. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$ as a function of time, cluster size indicated by colorbar. Default simulation parameters with labeled values of n_f.

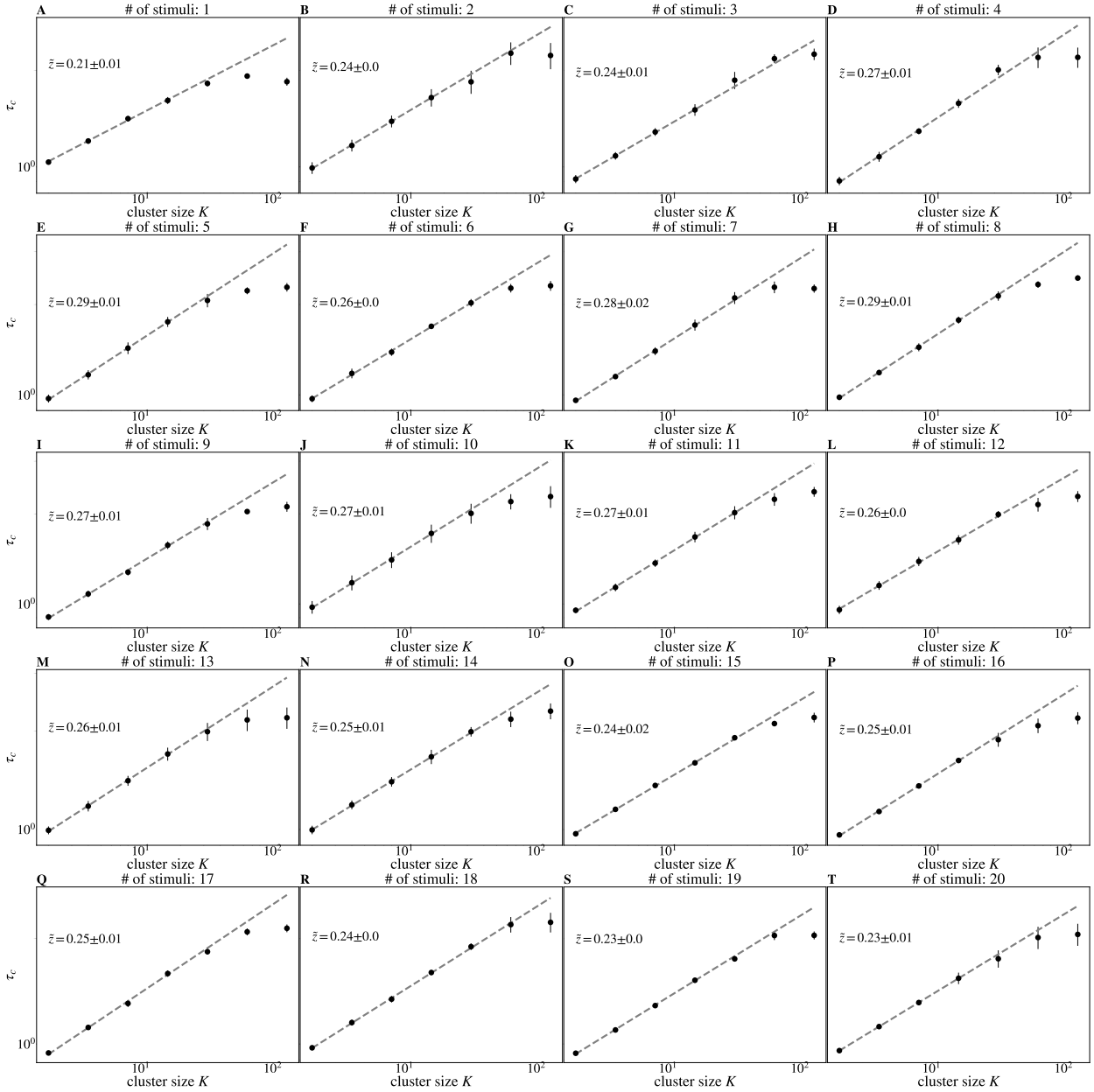

FIG. 41. Time constants τ_c extracted from each curve in **(A)**, and observe behavior obeying $\tau_c \propto K^{\tilde{z}}$ for roughly 1 decade. Default simulation parameters with labeled values of n_f.

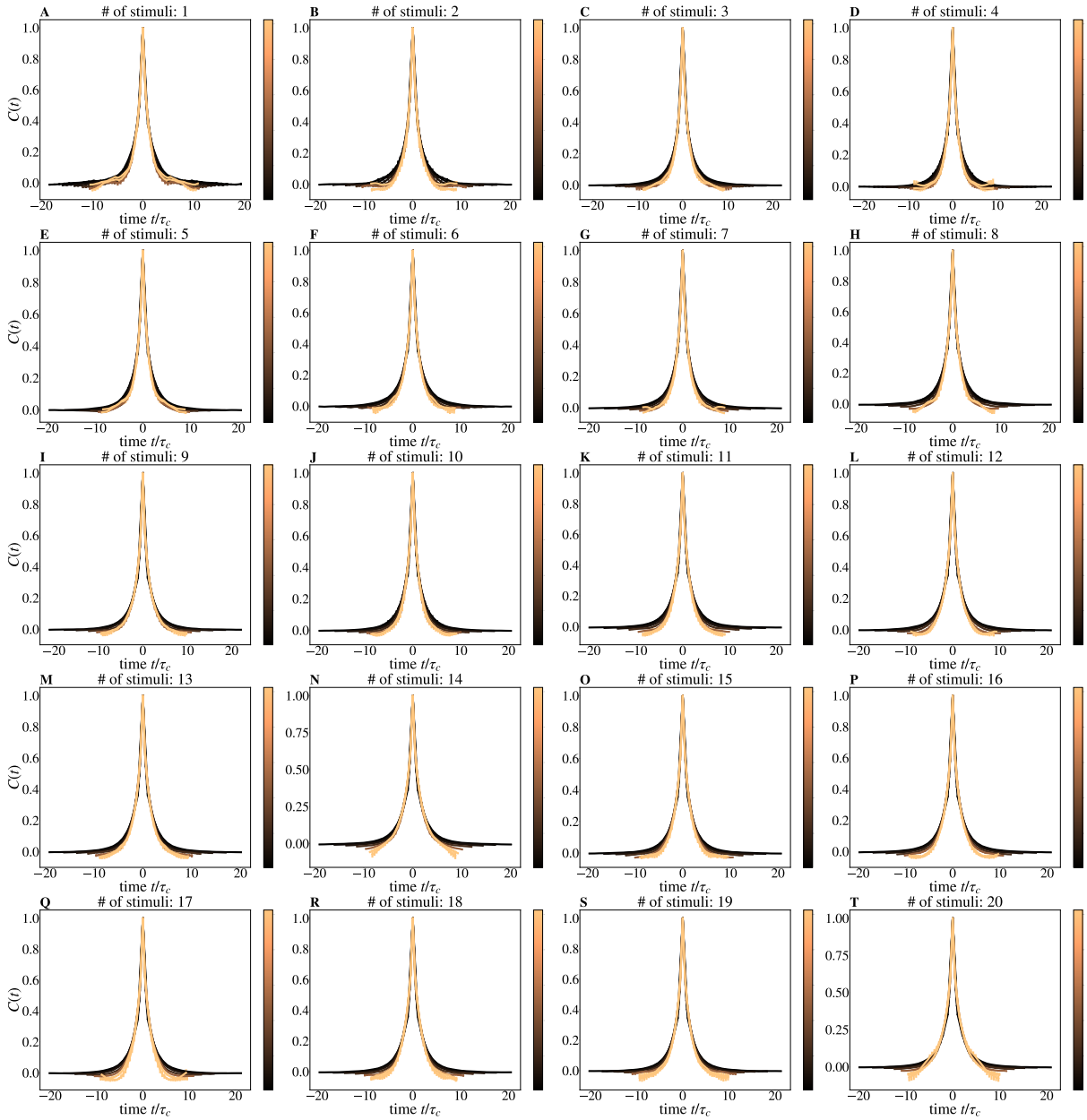

FIG. 42. Average autocorrelation function for cluster sizes $K = 2, 4, \ldots, 256$, where time is rescaled by the appropriate τ_c for that coarse-graining iteration. Default simulation parameters with labeled values of n_f.

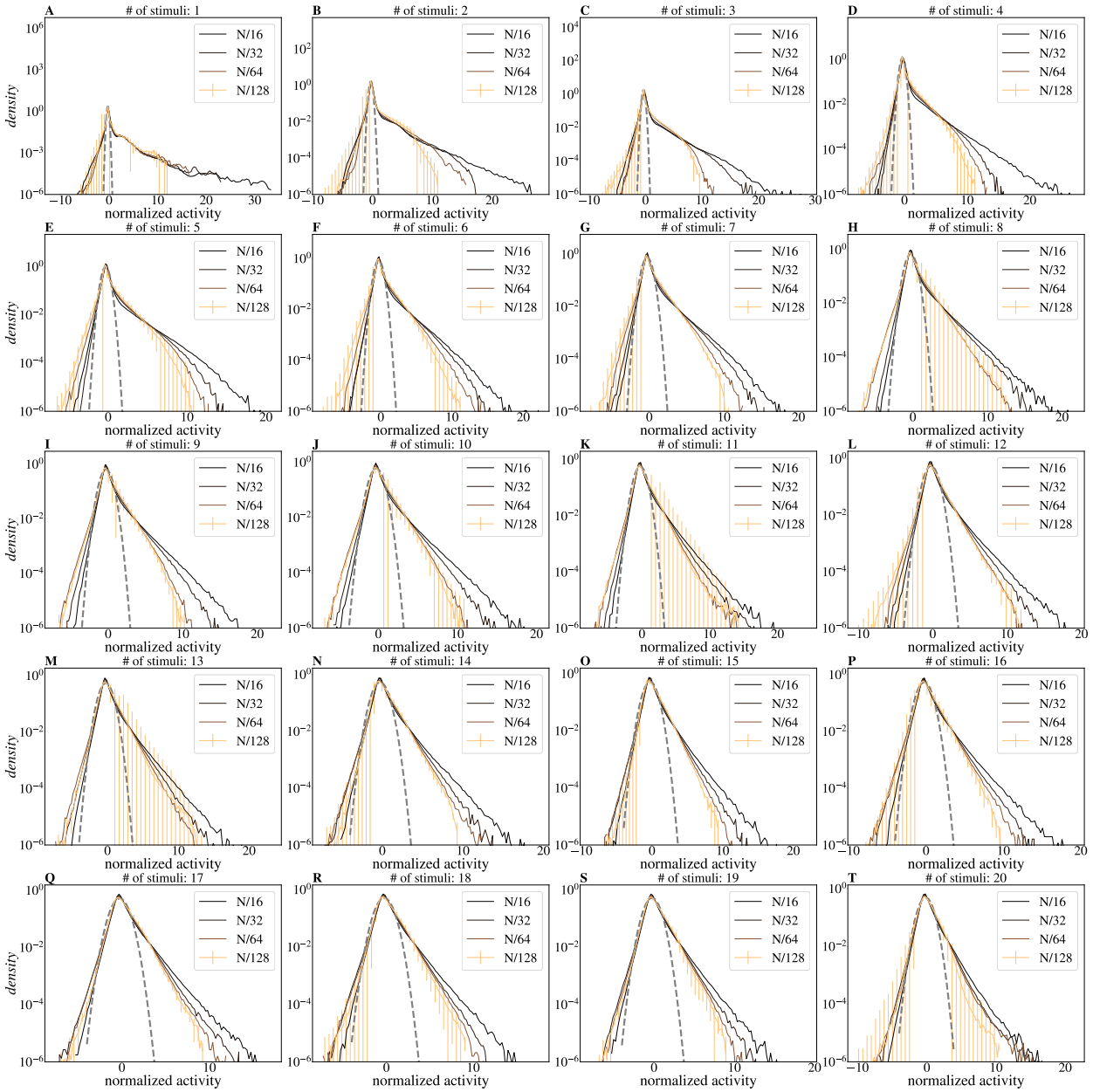

FIG. 43. Distribution of coarse-grained variables for $k = N/16, N/32, N/64, N/128$ modes retained. Default simulation parameters with labeled values of n_f. As n_f increases, distributions become increasingly short-tailed. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

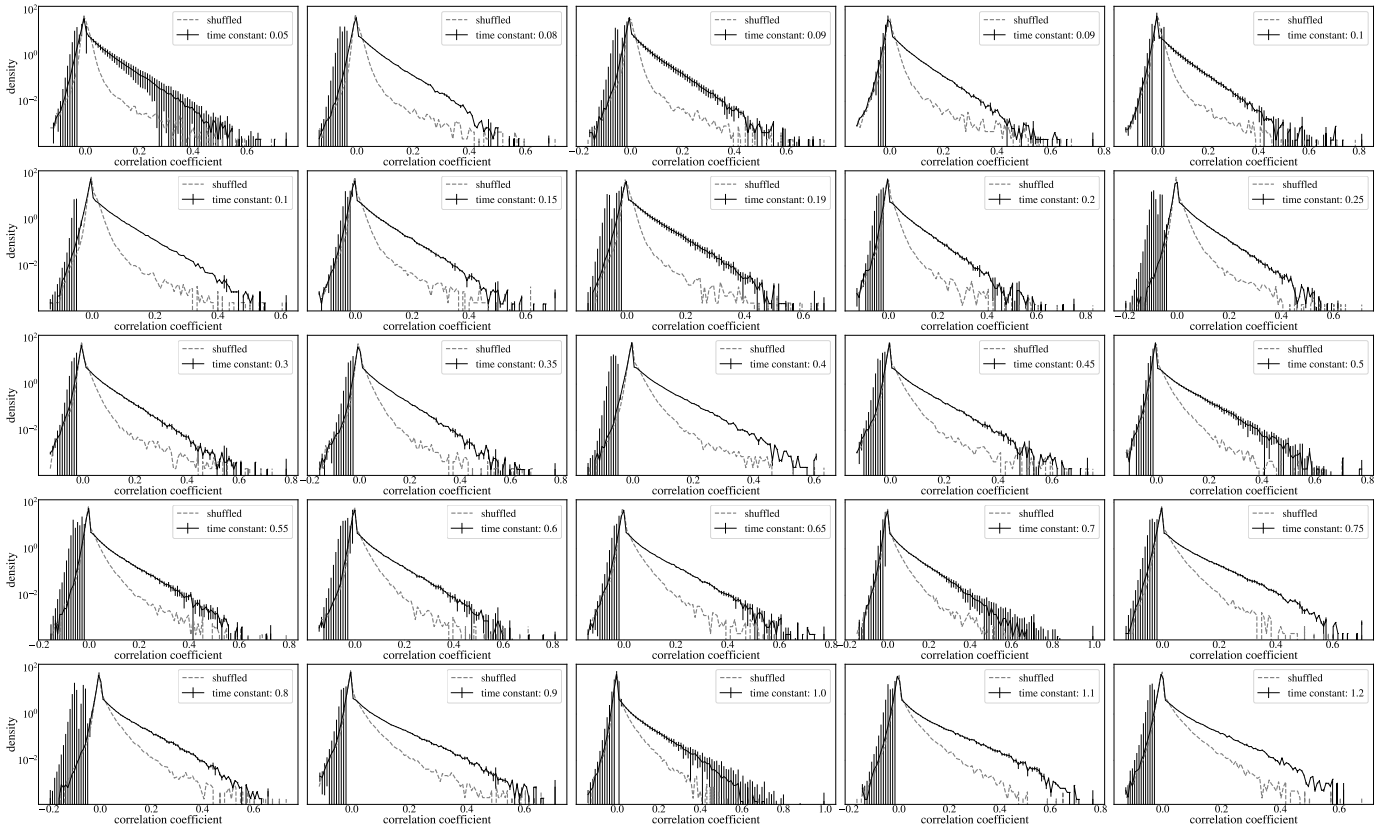

FIG. 44. Distributions of pairwise correlation coefficients. Default simulation parameters with value of τ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

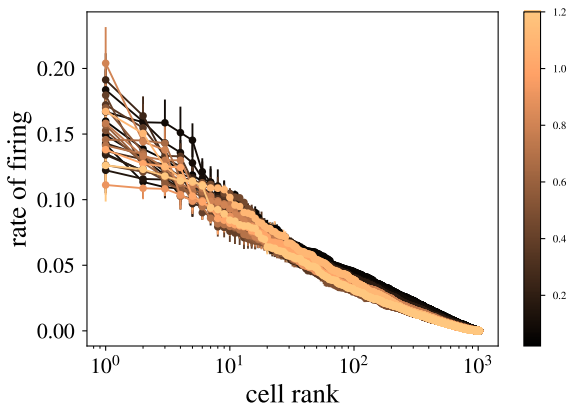

FIG. 45. Rate of firing vs rescaled rank. Colorbar indicates value of τ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

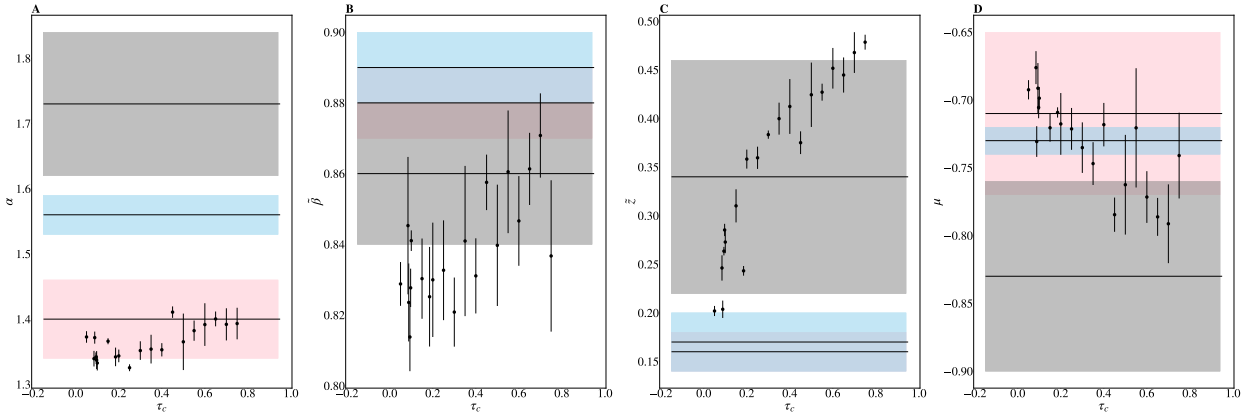

FIG. 46. Each critical exponent, $\alpha, \tilde{\beta}, \tilde{z}, \mu$ vs latent field time constant τ. Results from [1] marked and shaded in gray, pink, and blue. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

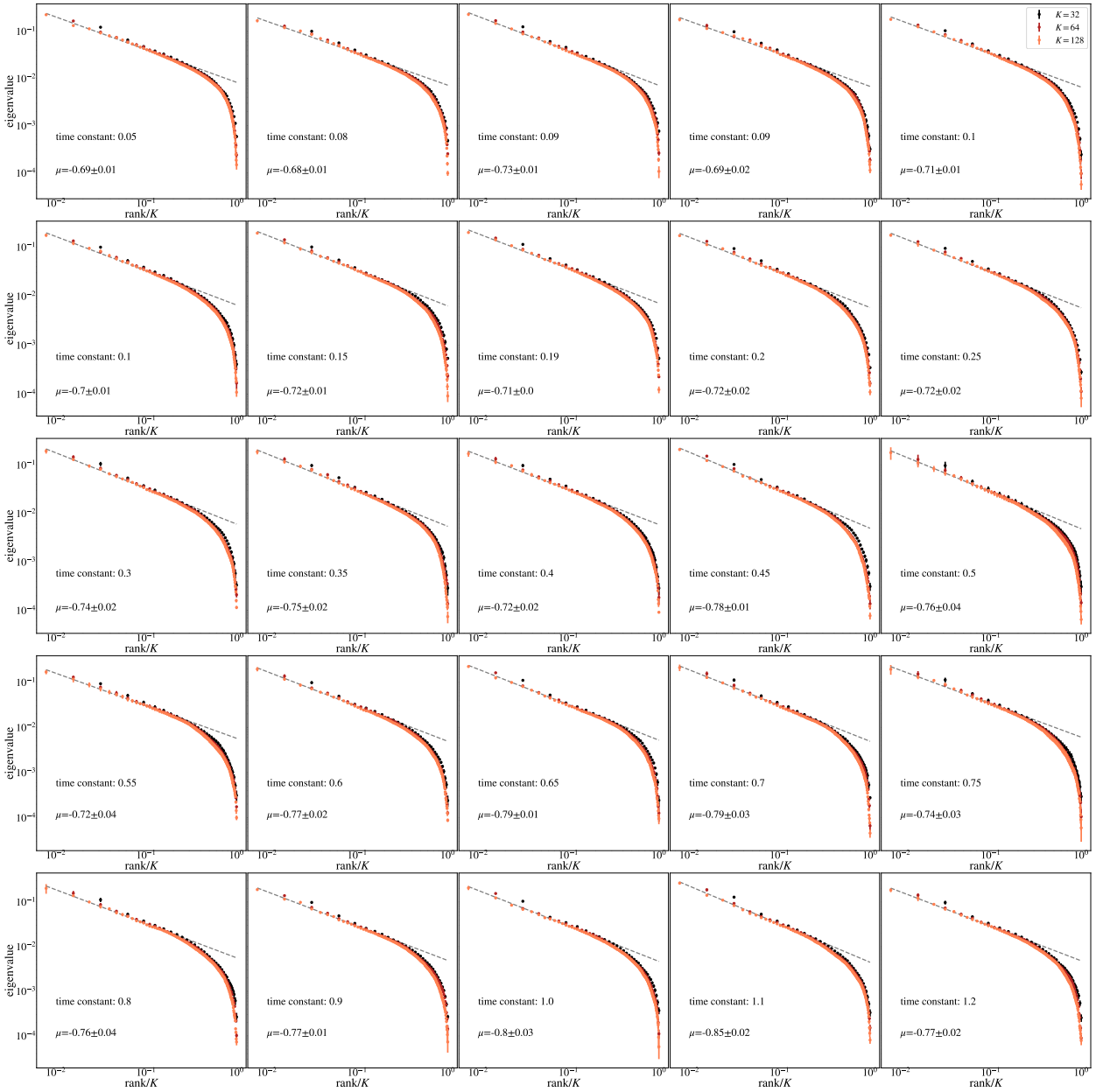

FIG. 47. Average eigenvalue spectrum of Eq 40 for cluster sizes $K = 32, 64, 128$. Default simulation parameters with labeled values of τ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

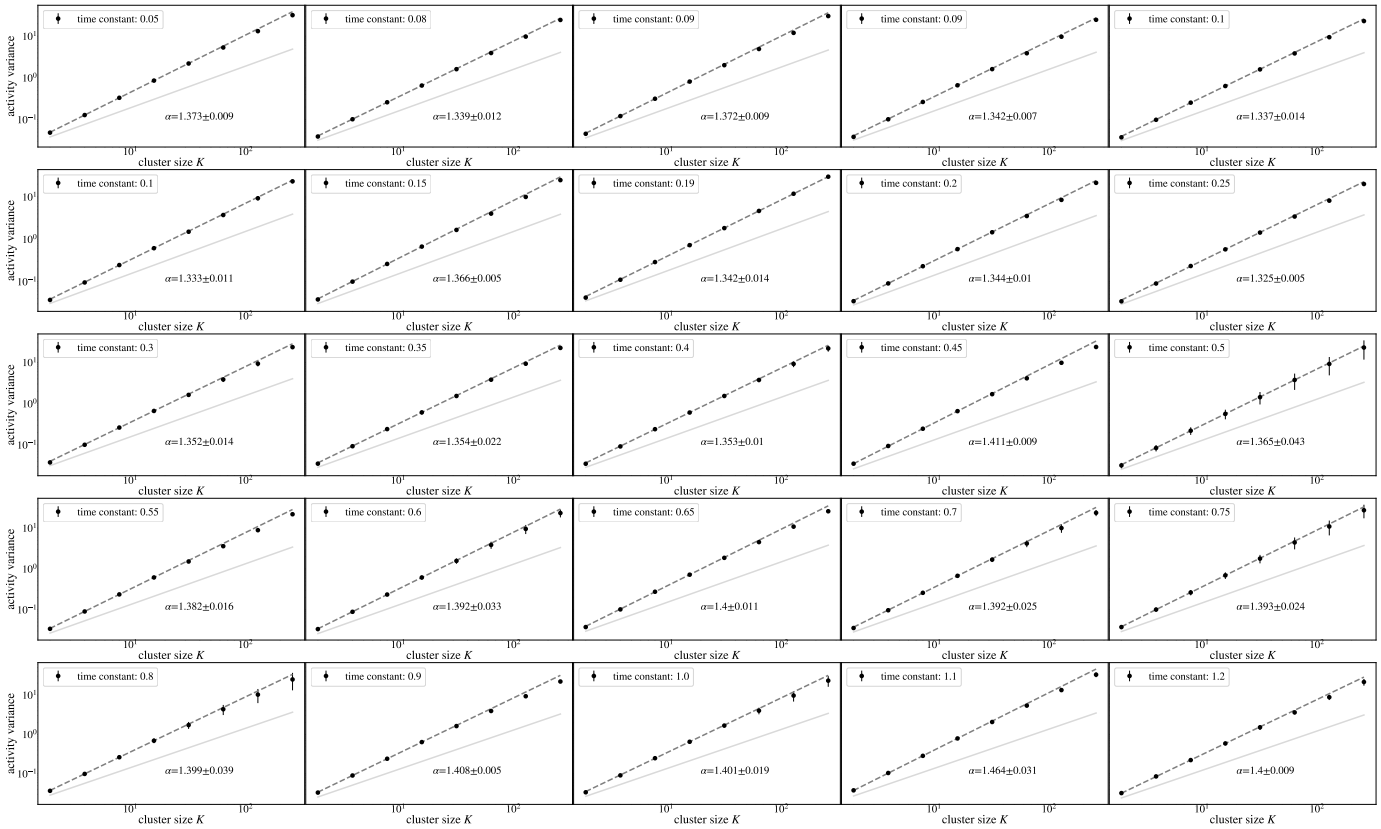

FIG. 48. Activity variance (Eq 33) over coarse-grained variables at each coarse-graining iteration. Default simulation parameters with labeled values of τ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

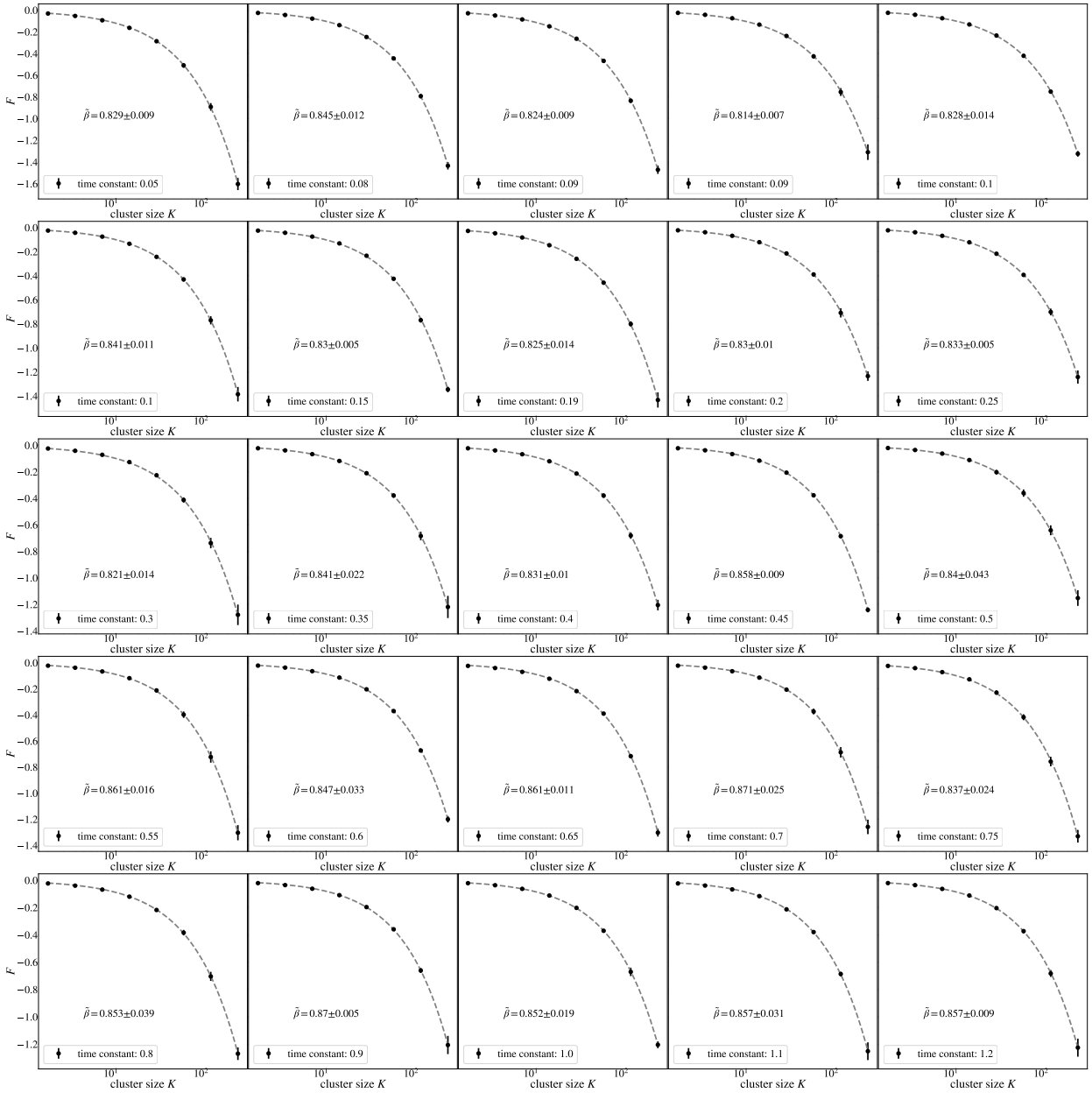

FIG. 49. Average free energy (Eq 34) at each coarse-graining iteration. Default simulation parameters with labeled values of τ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

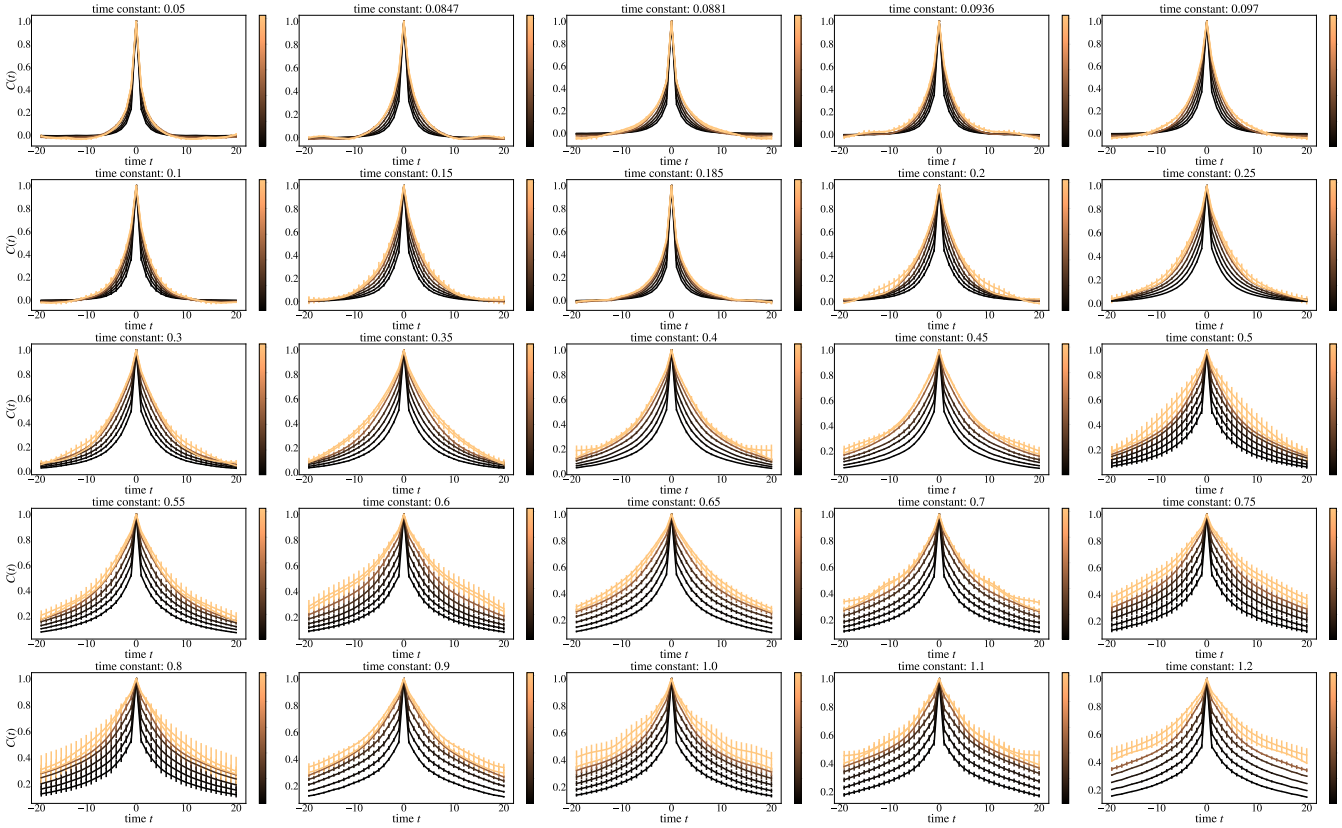

FIG. 50. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$ as a function of time, cluster size indicated by colorbar. Default simulation parameters with labeled values of τ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

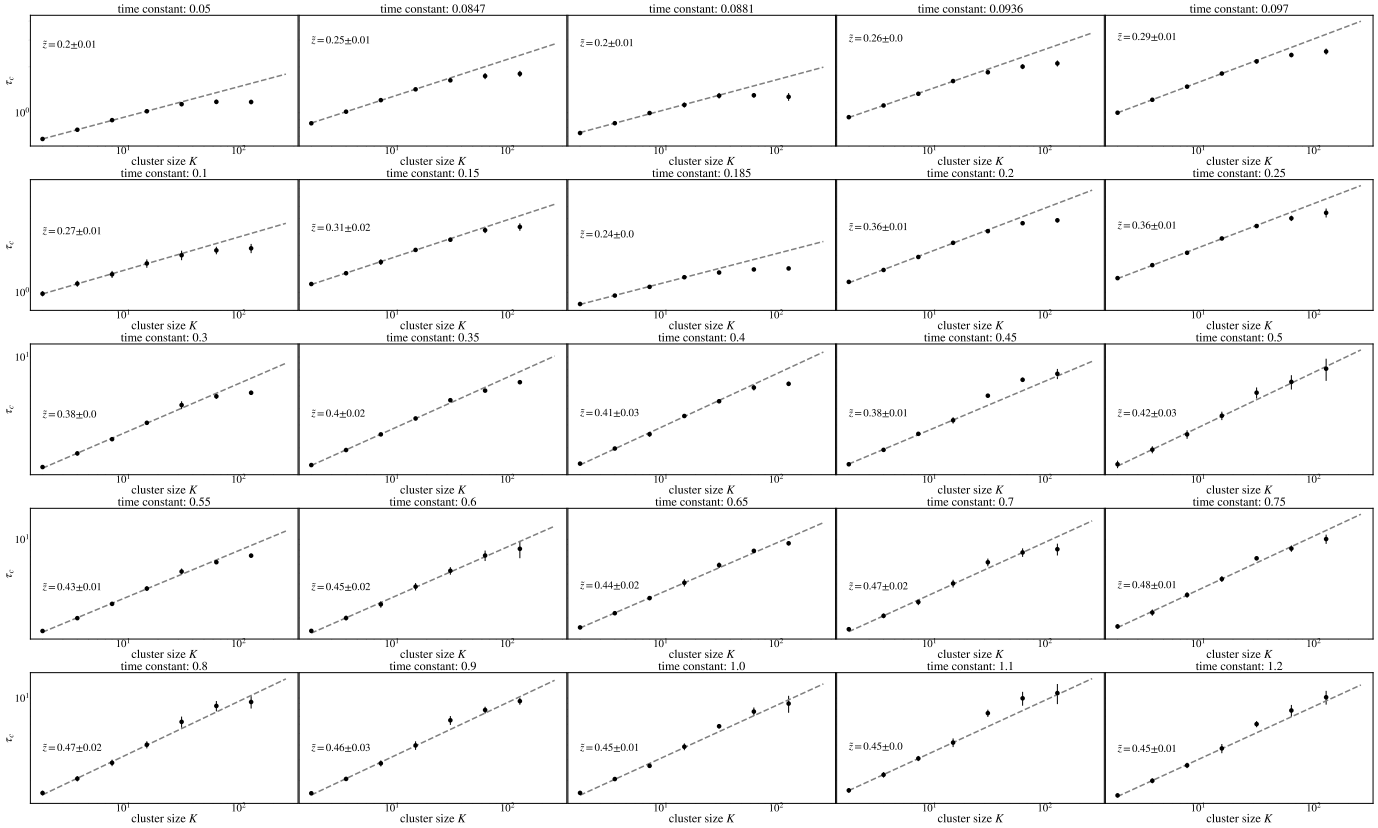

FIG. 51. Time constants τ_c extracted from each curve in in FIG 50, and observe behavior obeying $\tau_c \propto K^{\tilde{z}}$ for roughly 1 decade. Default simulation parameters with labeled values of τ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

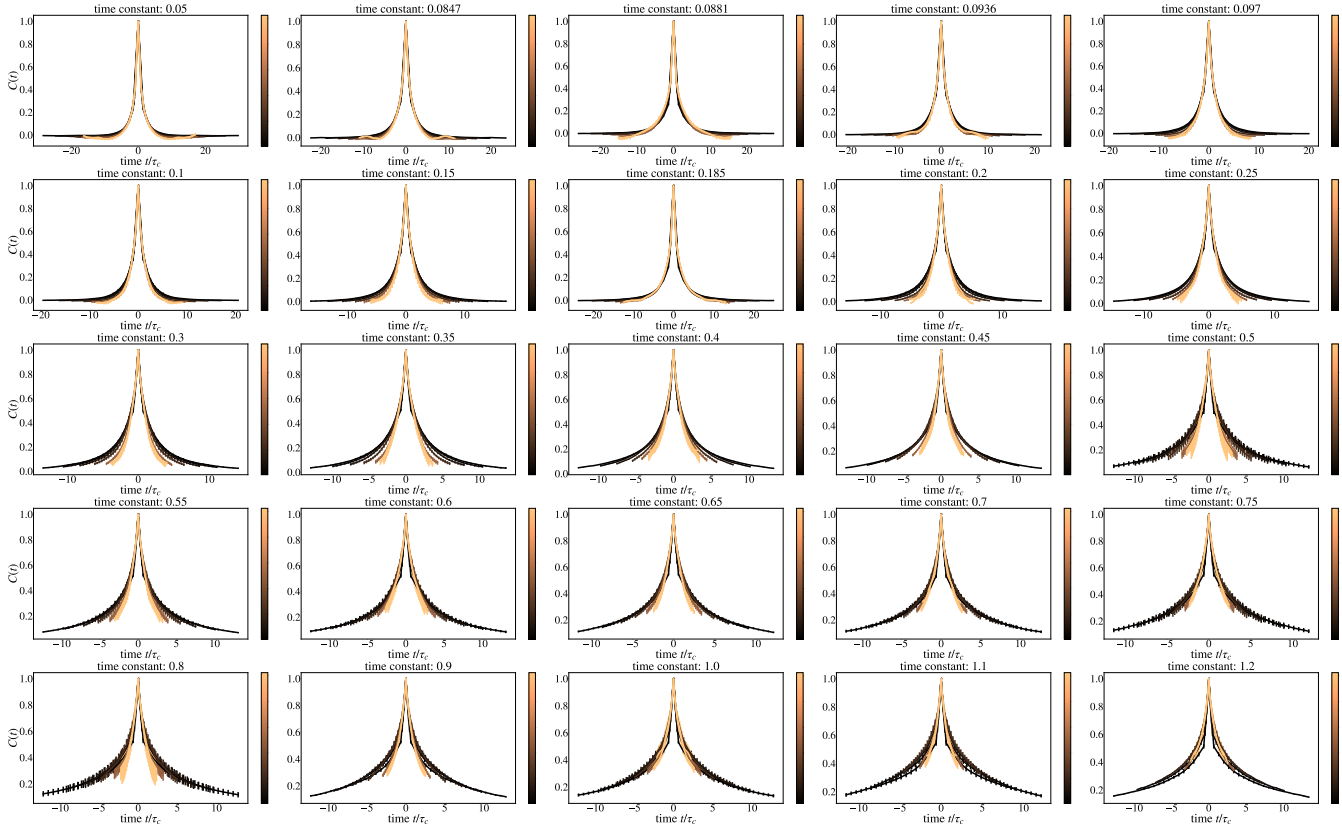

FIG. 52. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$, where time is rescaled by the appropriate τ_c for that coarse-graining iteration. Default simulation parameters with labeled values of τ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

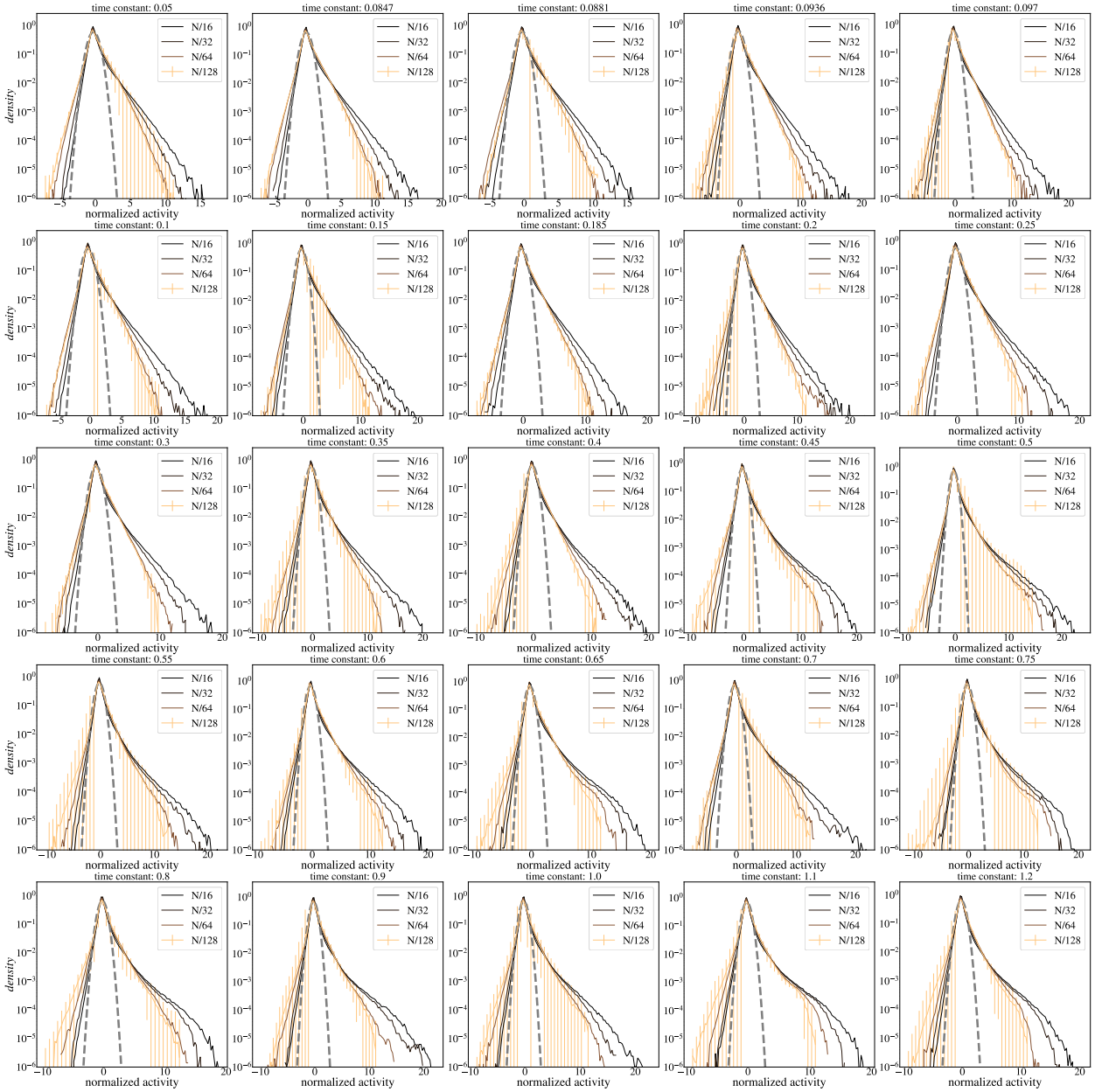

FIG. 53. Distribution of coarse-grained variables for $k = N/16, N/32, N/64, N/128$ modes retained. Default simulation parameters with labeled values of τ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

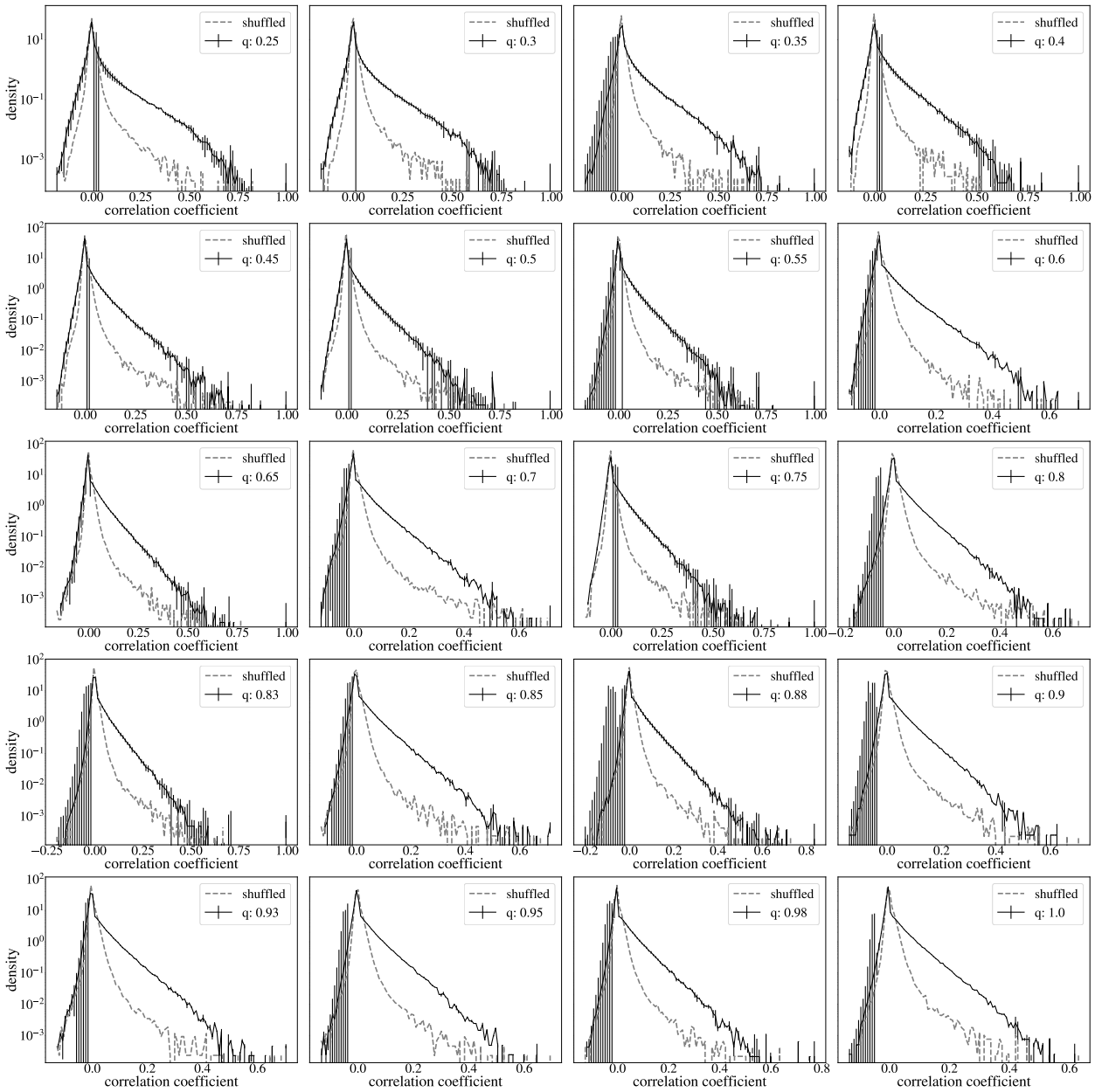

FIG. 54. Distributions of pairwise correlation coefficients. Default simulation parameters with value of q. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

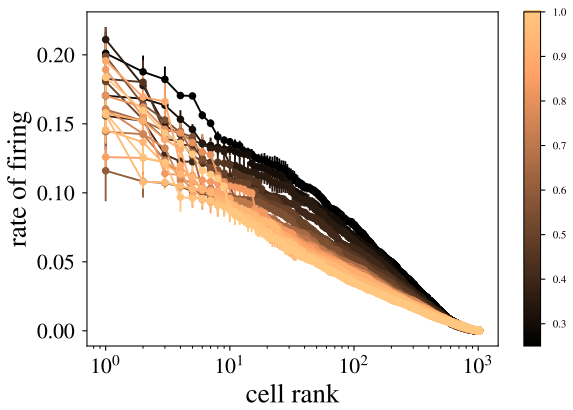

FIG. 55. Rate of firing vs rescaled rank. Colorbar indicates value of q

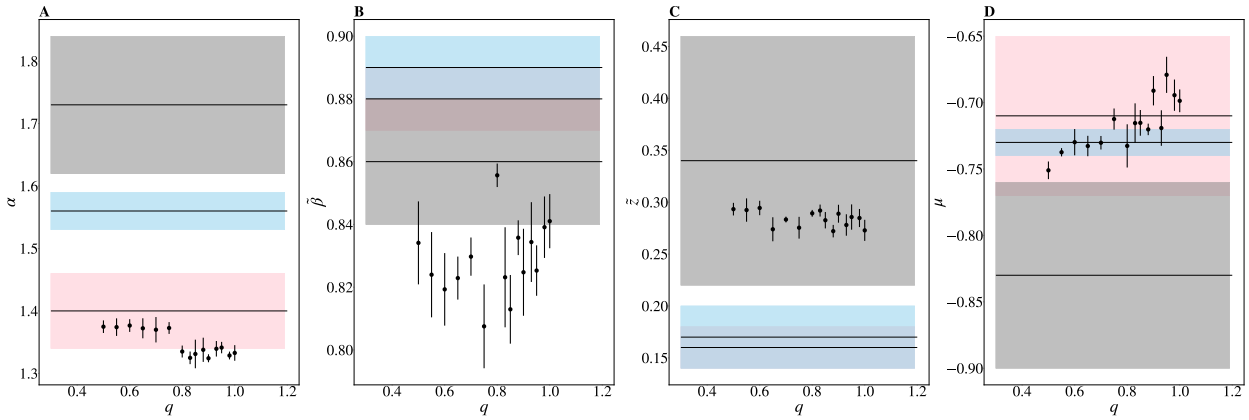

FIG. 56. Each critical exponent, $\alpha, \tilde{\beta}, \tilde{z}, \mu$ vs probability of coupling to a latent field q. Results from [1] marked and shaded in gray, pink, and blue. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

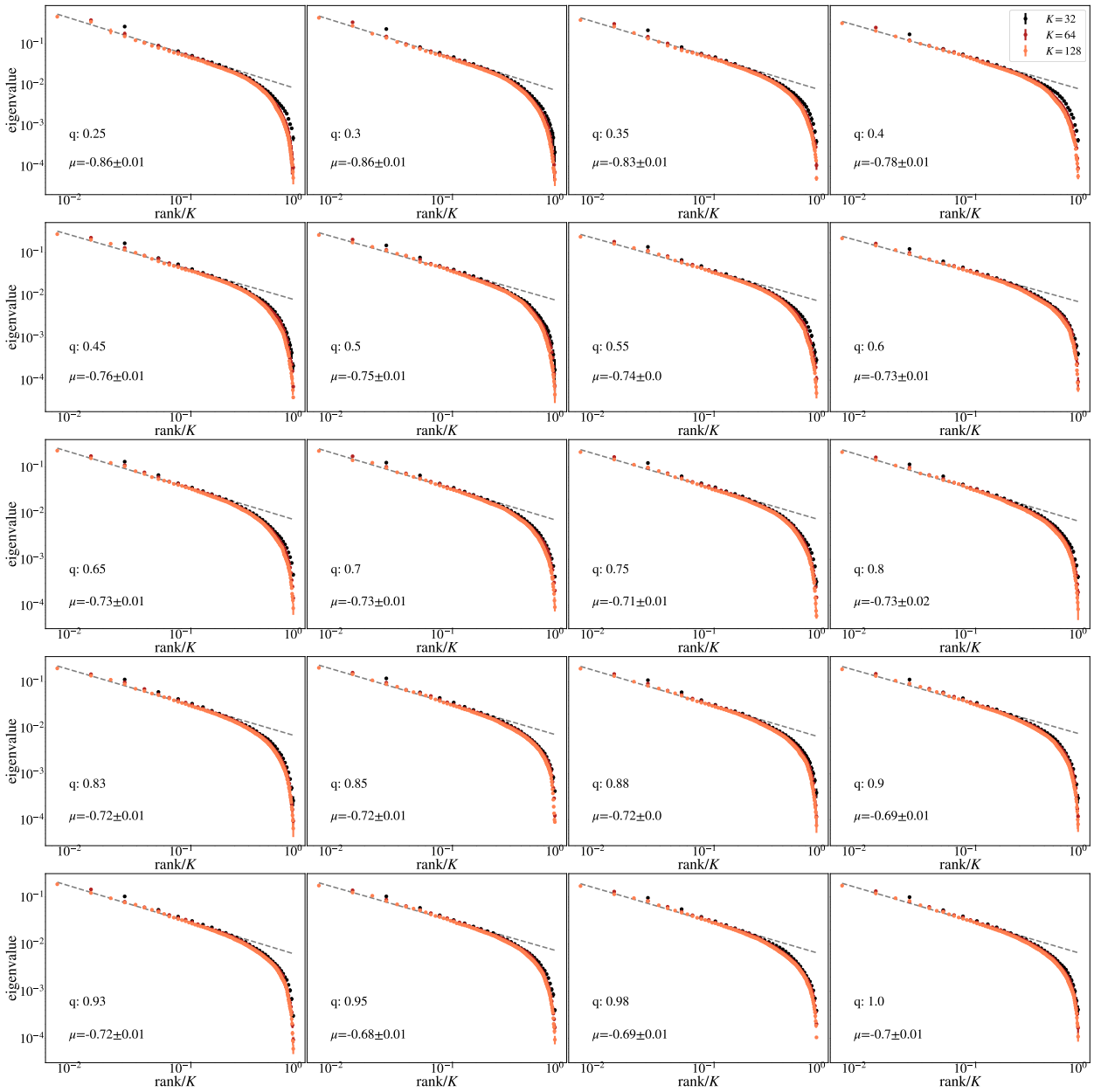

FIG. 57. Average eigenvalue spectrum of Eq 40 for cluster sizes $K = 32, 64, 128$. Default simulation parameters with labeled values of q. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

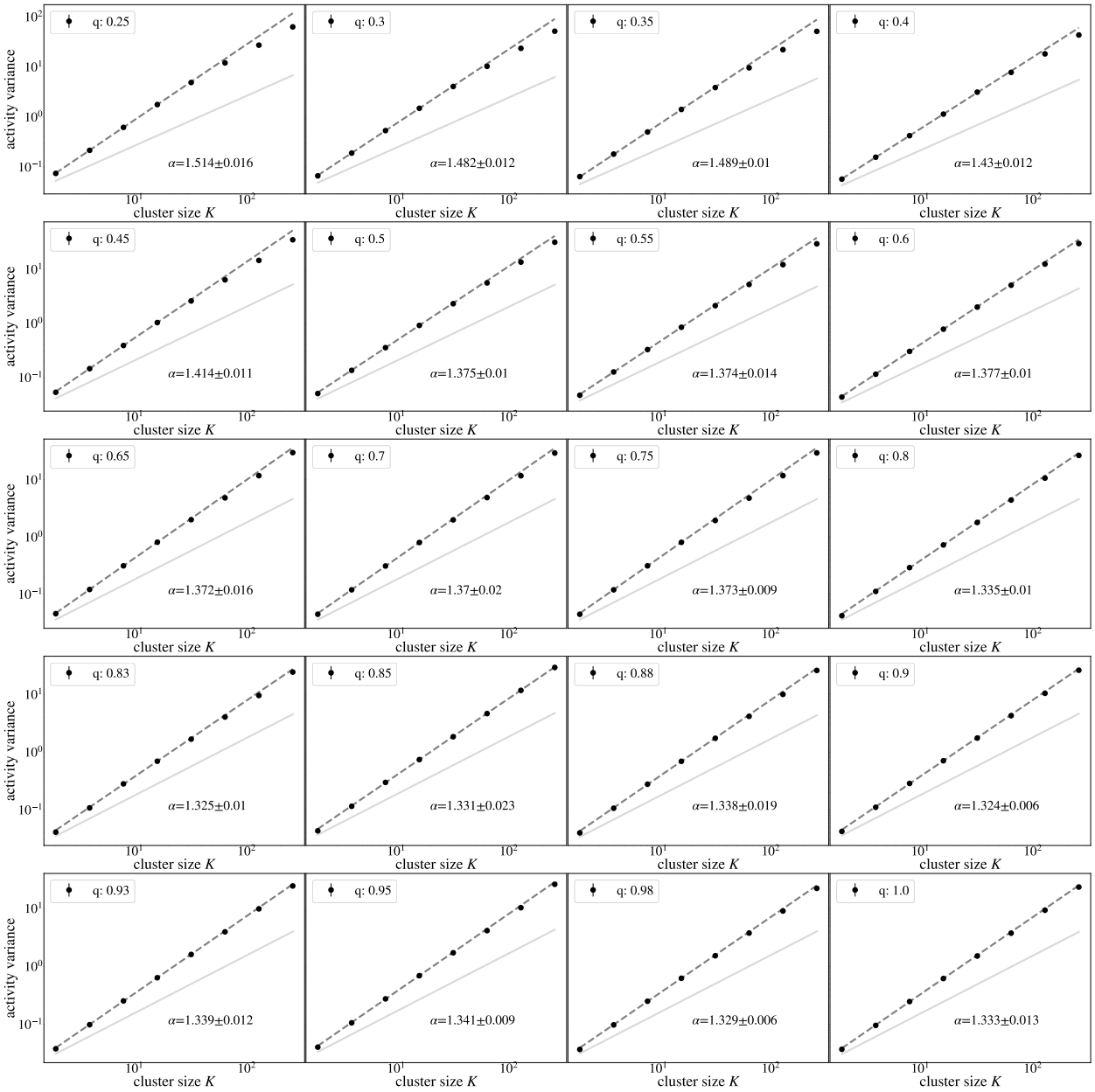

FIG. 58. Activity variance (Eq 33) over coarse-grained variables at each coarse-graining iteration. Default simulation parameters with labeled values of q. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

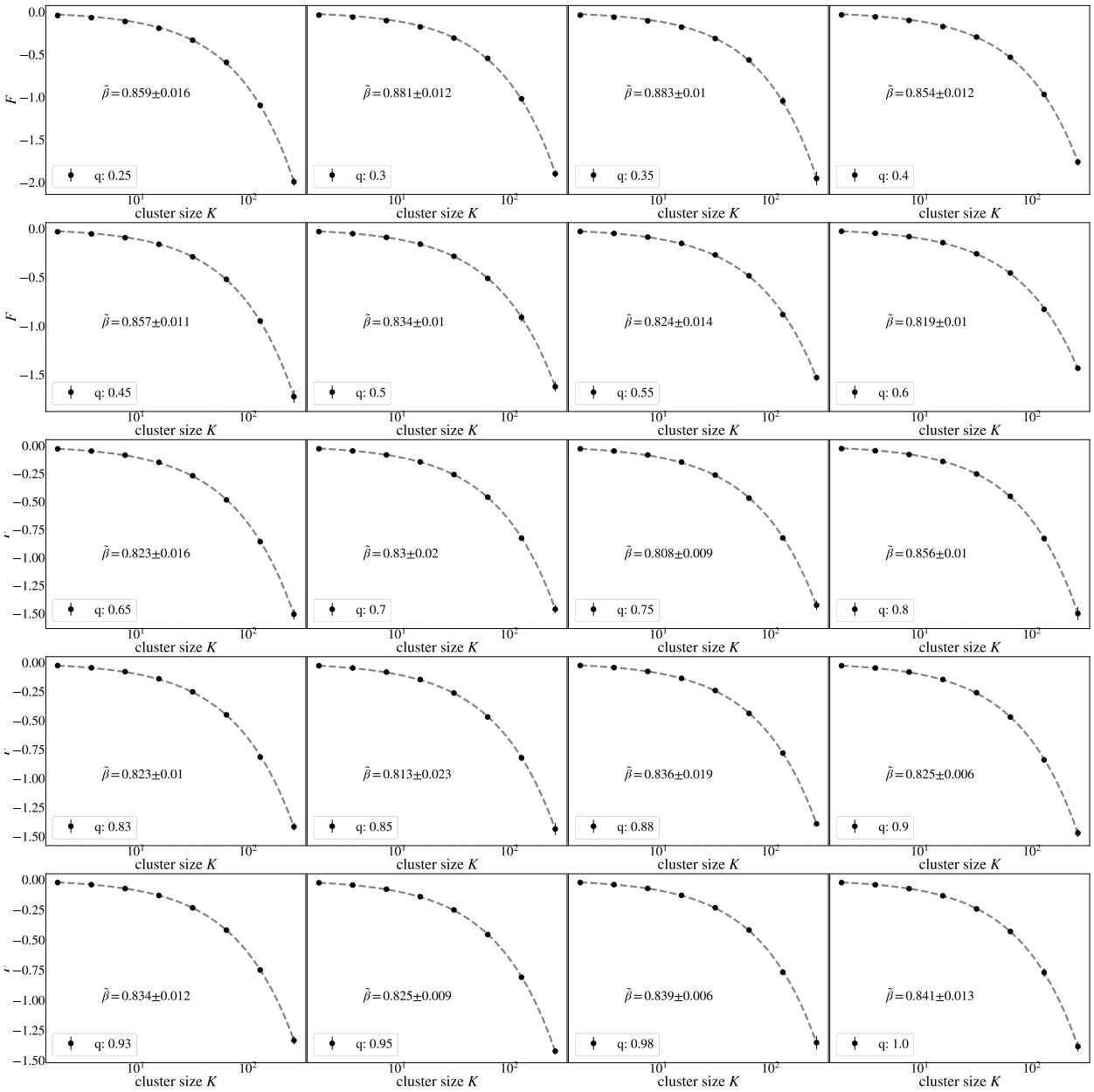

FIG. 59. Average free energy (Eq 34) at each coarse-graining iteration. Default simulation parameters with labeled values of q. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

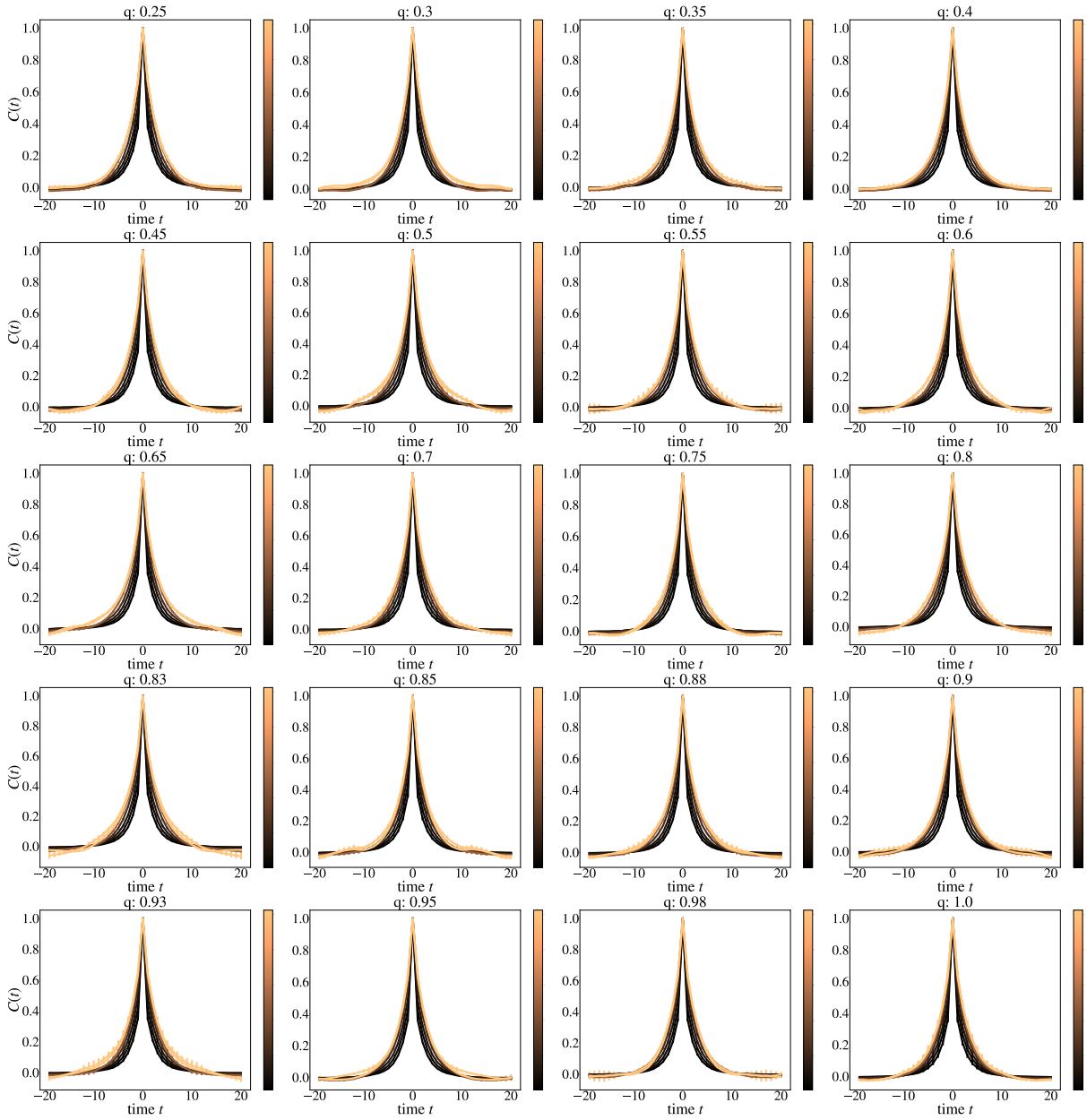

FIG. 60. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$ as a function of time, cluster size indicated by colorbar. Default simulation parameters with labeled values of q. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

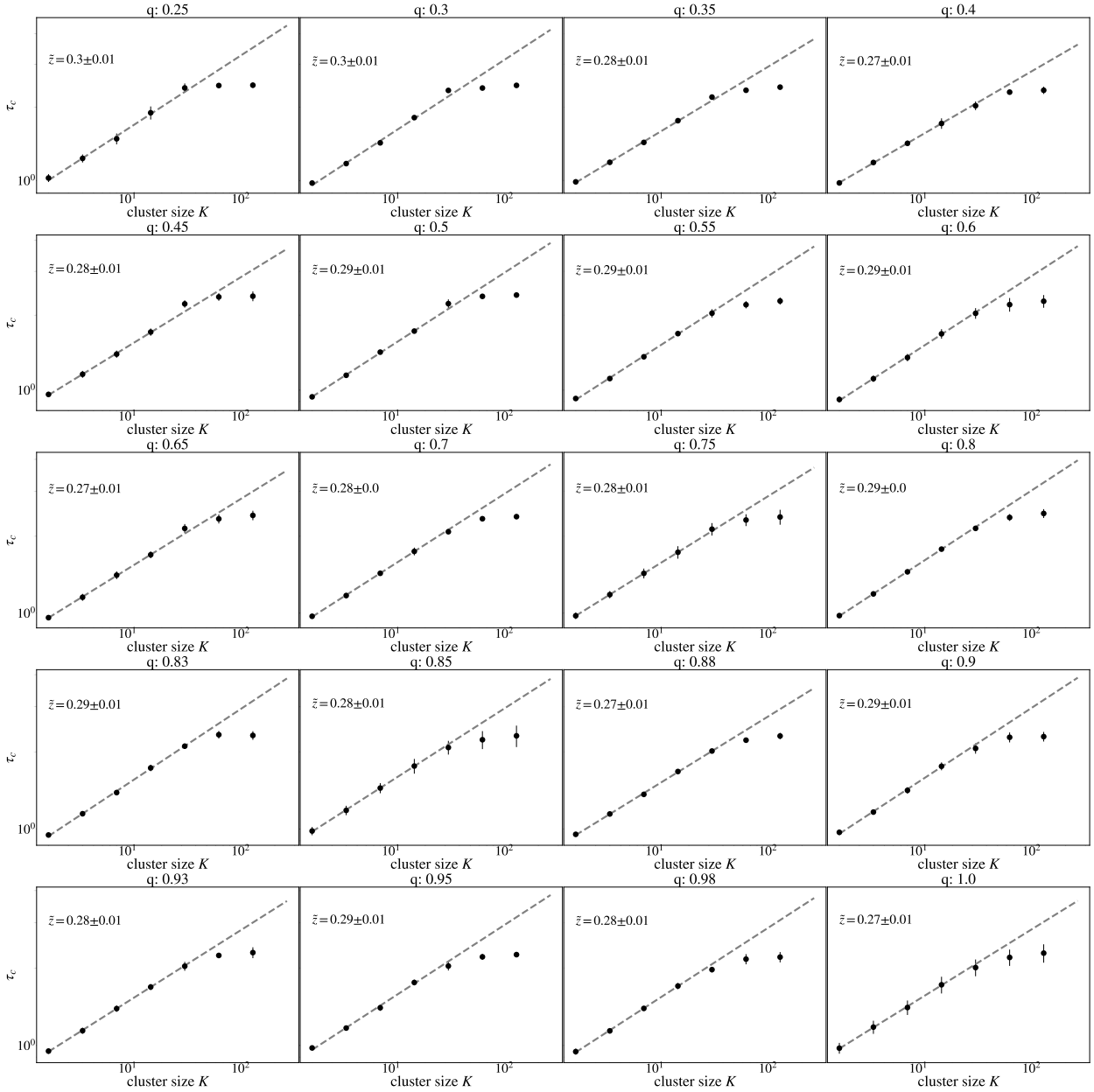

FIG. 61. Time constants τ_c extracted from each curve in in FIG 60, and observe behavior obeying $\tau_c \propto K^{\tilde{z}}$ for roughly 1 decade. Default simulation parameters with labeled values of q. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

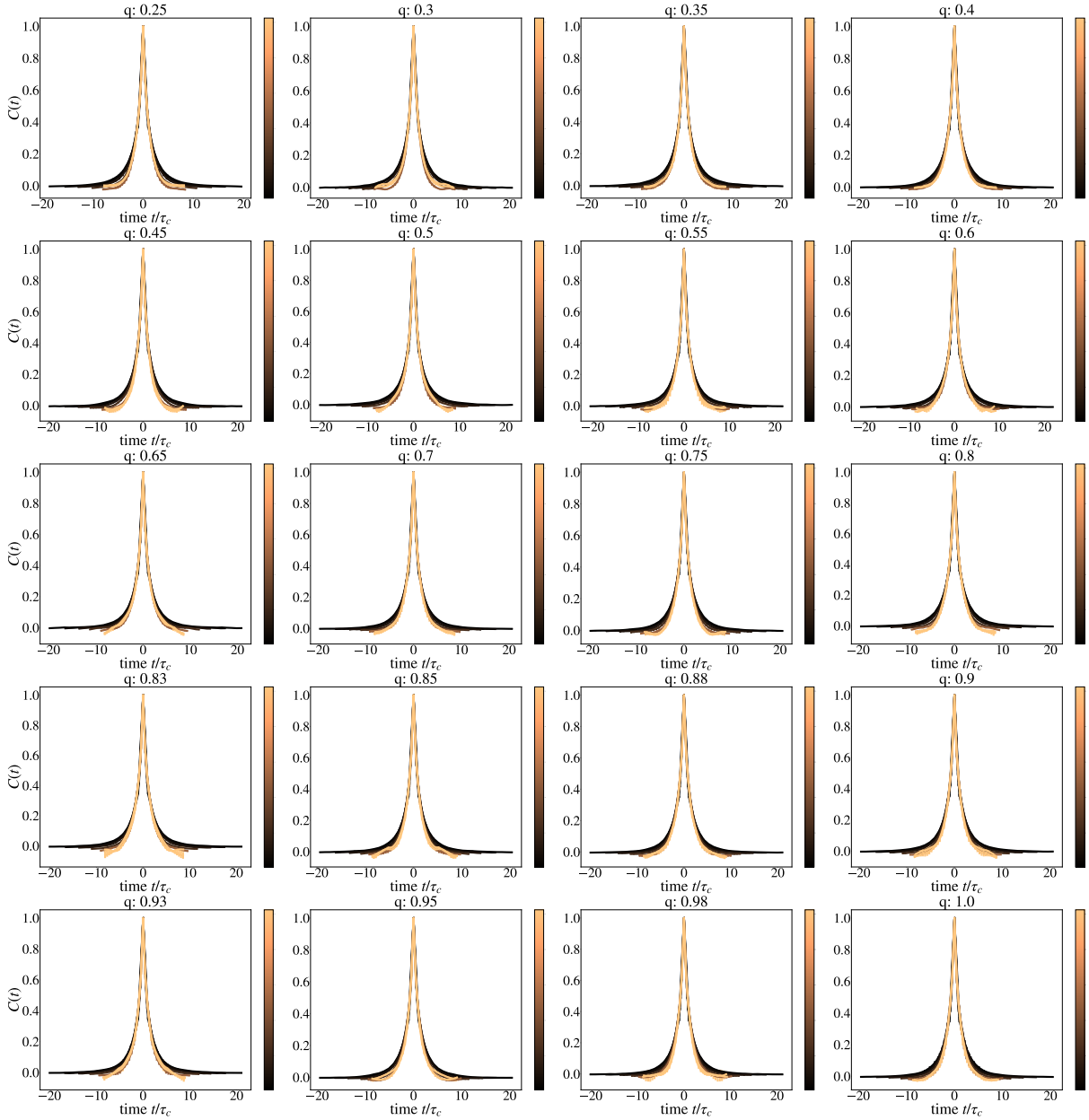

FIG. 62. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$, where time is rescaled by the appropriate τ_c for that coarse-graining iteration. Default simulation parameters with labeled values of q. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

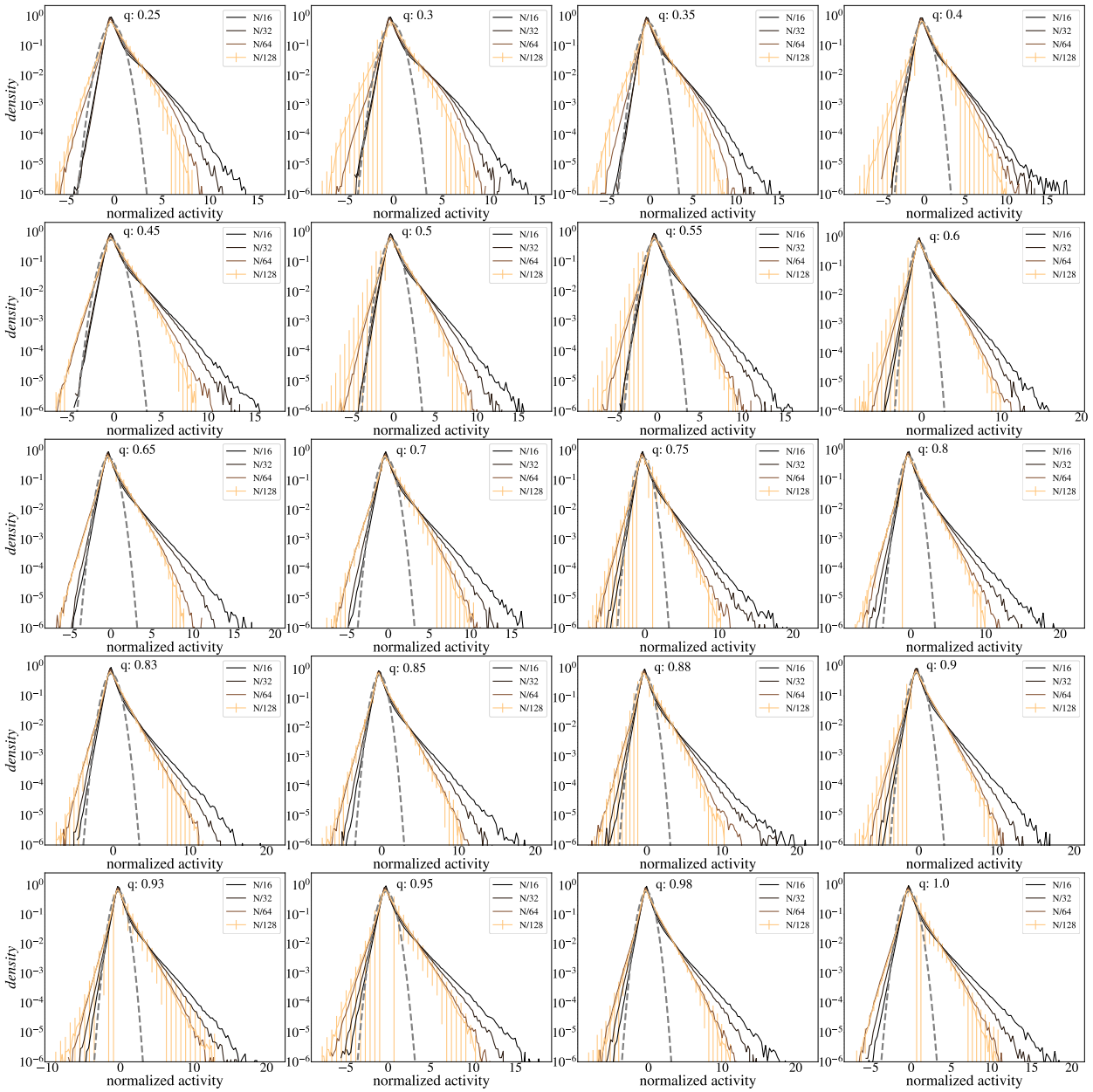

FIG. 63. Distribution of coarse-grained variables for $k = N/16, N/32, N/64, N/128$ modes retained. Default simulation parameters with labeled values of q. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

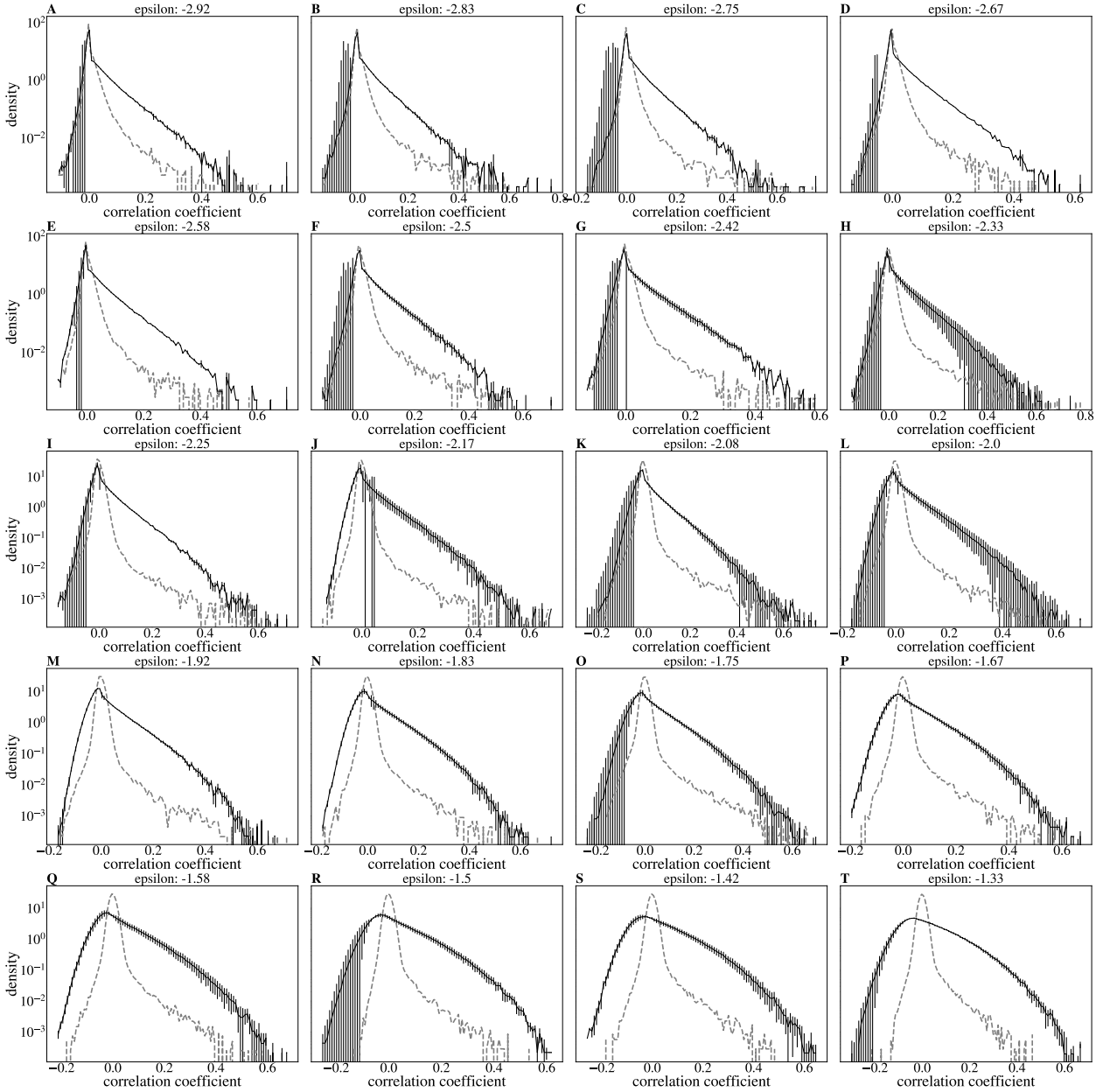

FIG. 64. Distributions of pairwise correlation coefficients. Default simulation parameters with value of ε. Error bars are standard deviations over randomly selected contiguous quarters of the simulation

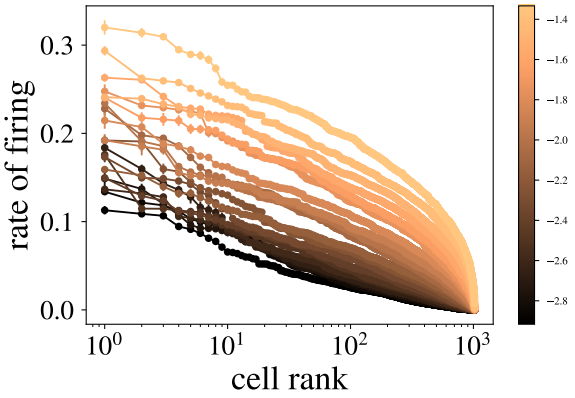

FIG. 65. Rate of firing vs rescaled rank. Colorbar indicates value of ε

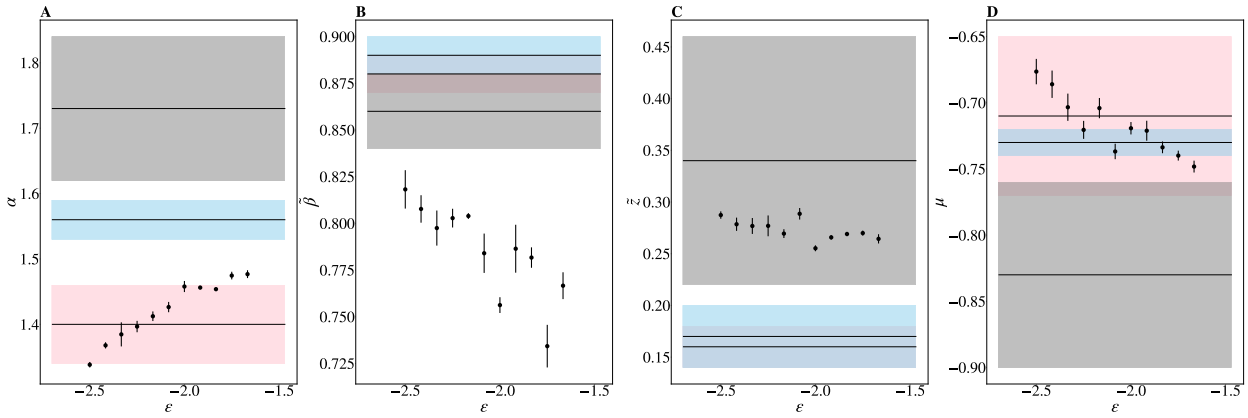

FIG. 66. Each critical exponent, $\alpha, \tilde{\beta}, \tilde{z}, \mu$ vs penalty term ε. Results from [1] marked and shaded in gray, pink, and blue. Error bars are standard deviations over randomly selected contiguous quarters of the simulation

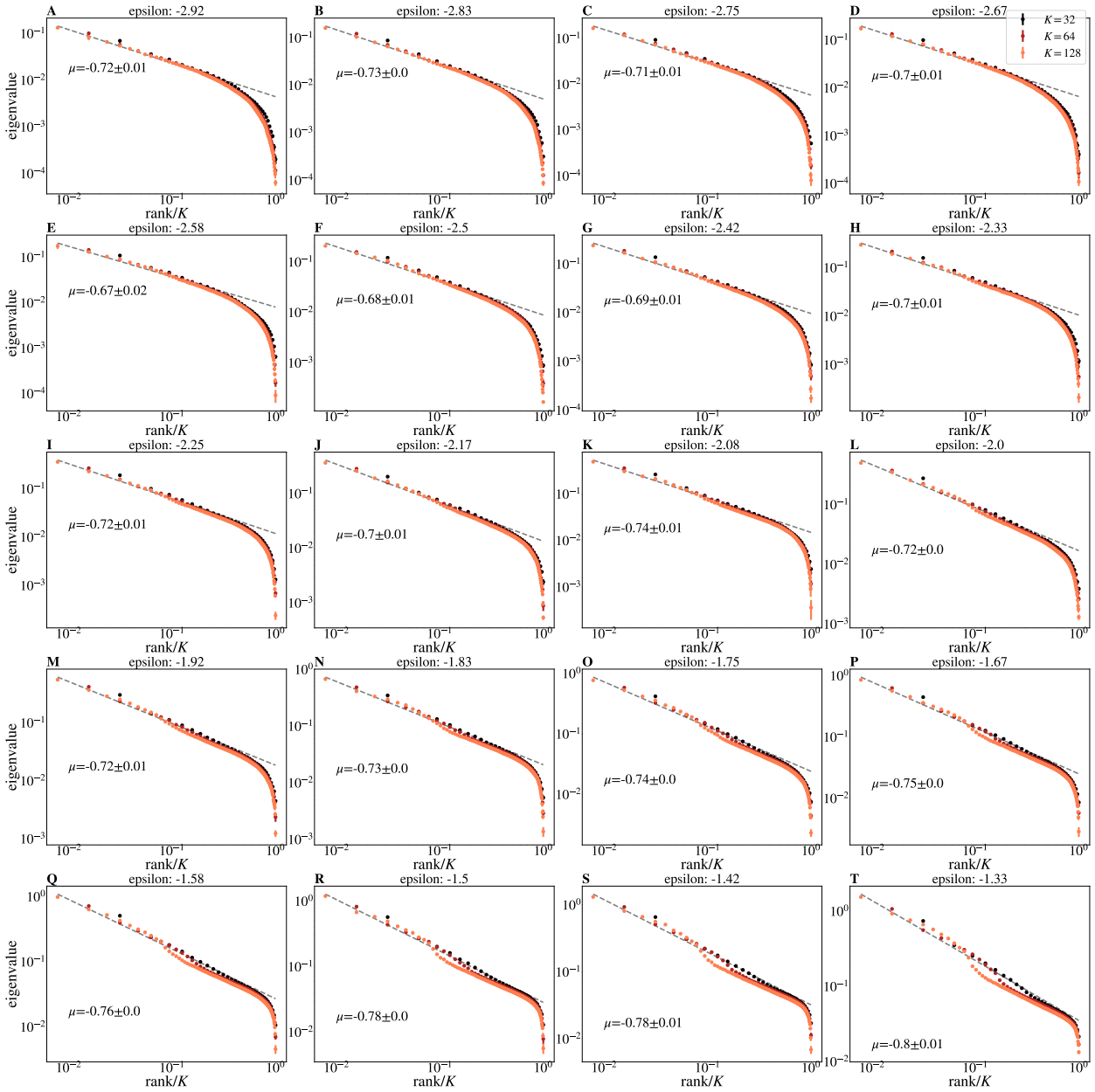

FIG. 67. Average eigenvalue spectrum of Eq 40 for cluster sizes $K = 32, 64, 128$. Default simulation parameters with labeled values of ε. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

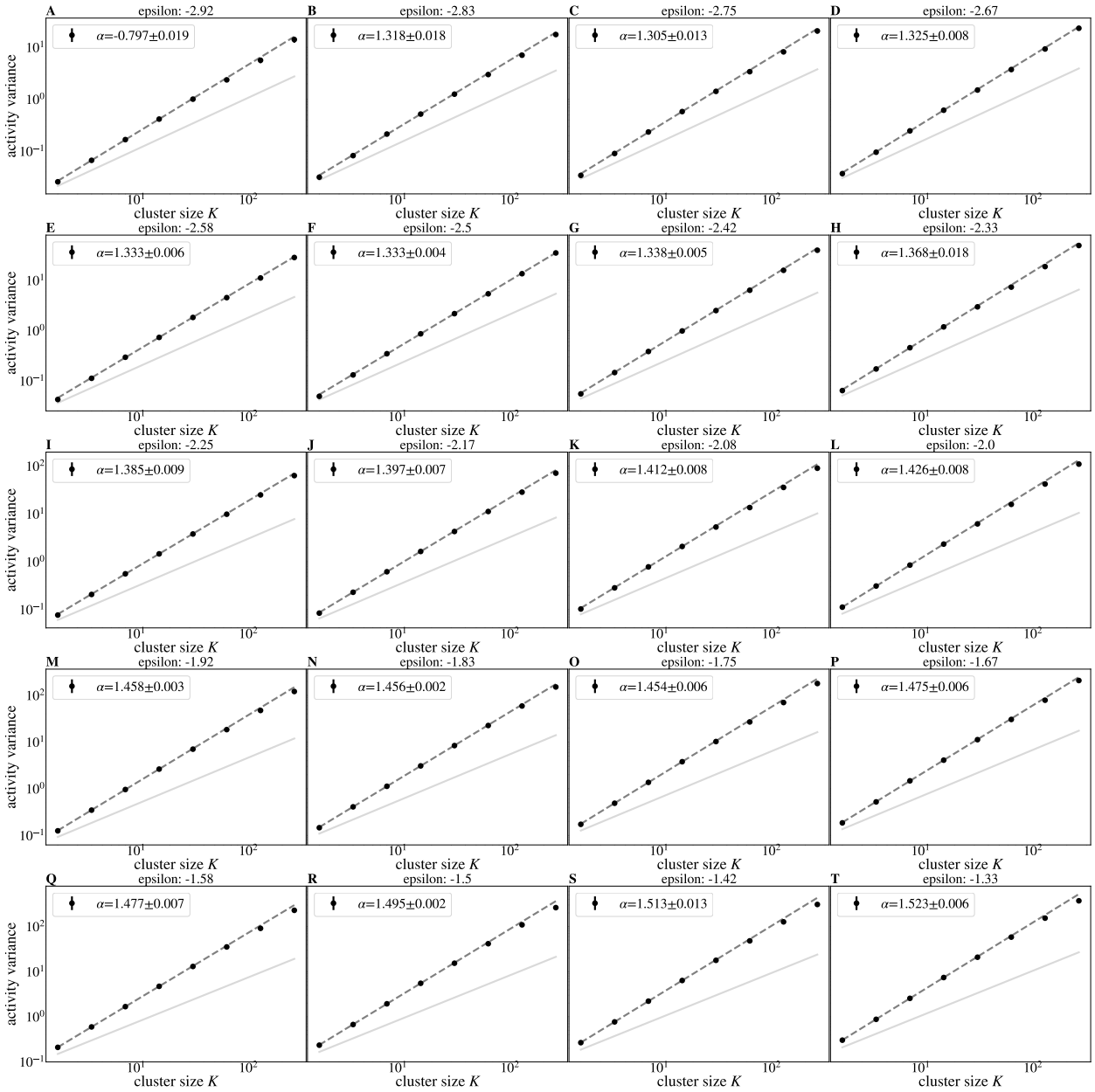

FIG. 68. Activity variance (Eq 33) over coarse-grained variables at each coarse-graining iteration. Default simulation parameters with labeled values of ε. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

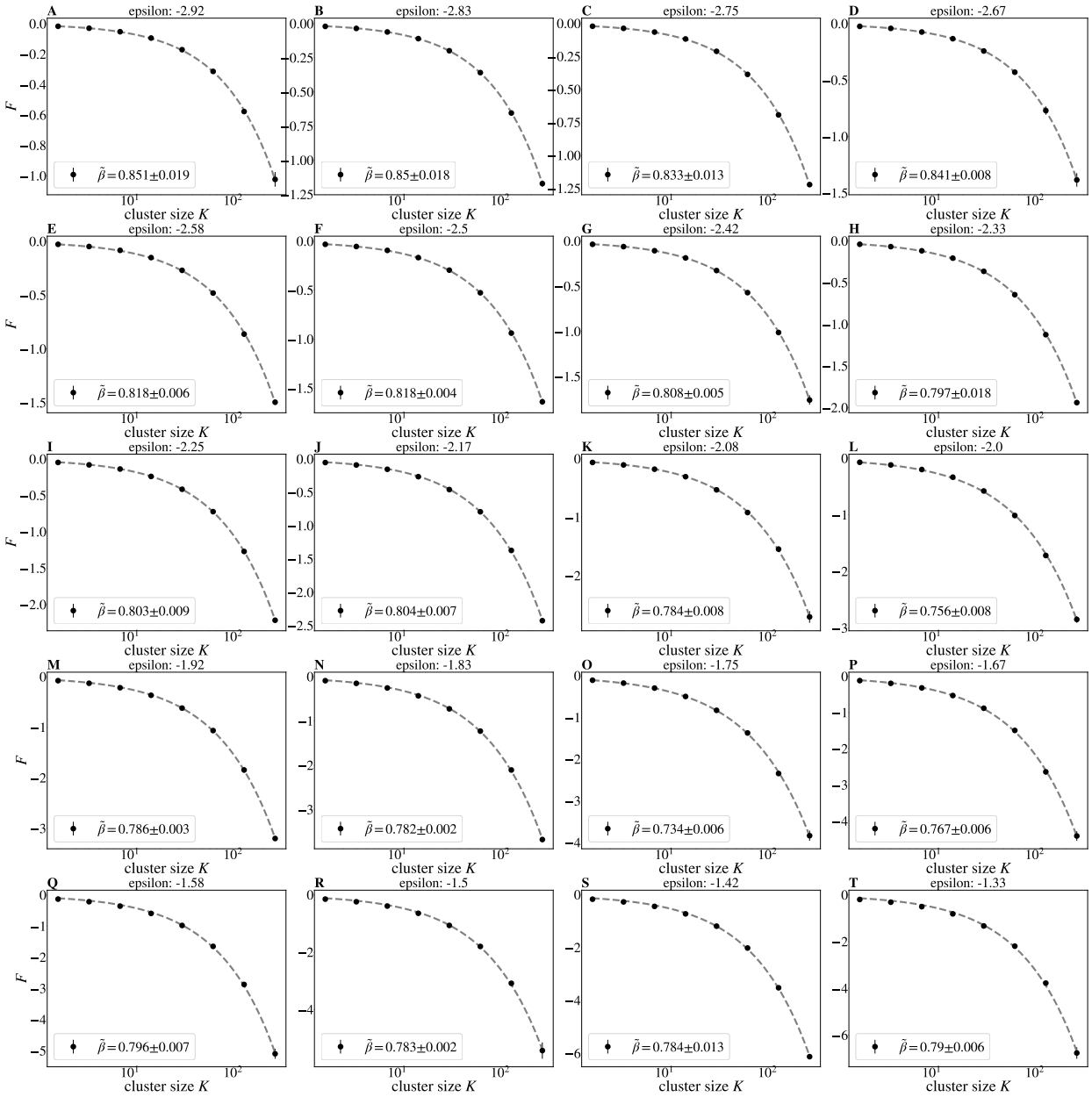

FIG. 69. Average free energy (Eq 34) at each coarse-graining iteration. Default simulation parameters with labeled values of ε.Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

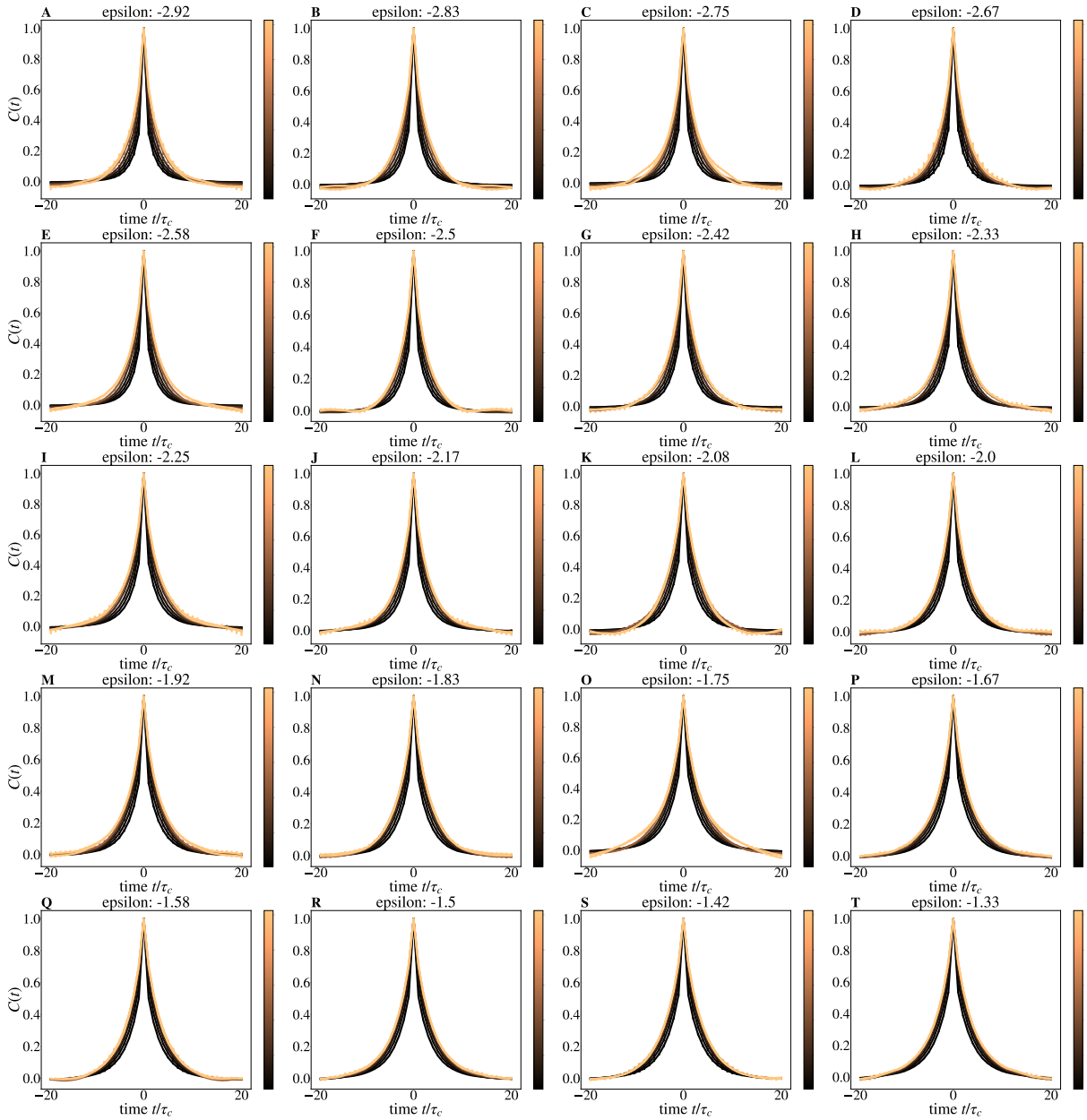

FIG. 70. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$ as a function of time, cluster size indicated by colorbar. Default simulation parameters with labeled values of ε. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

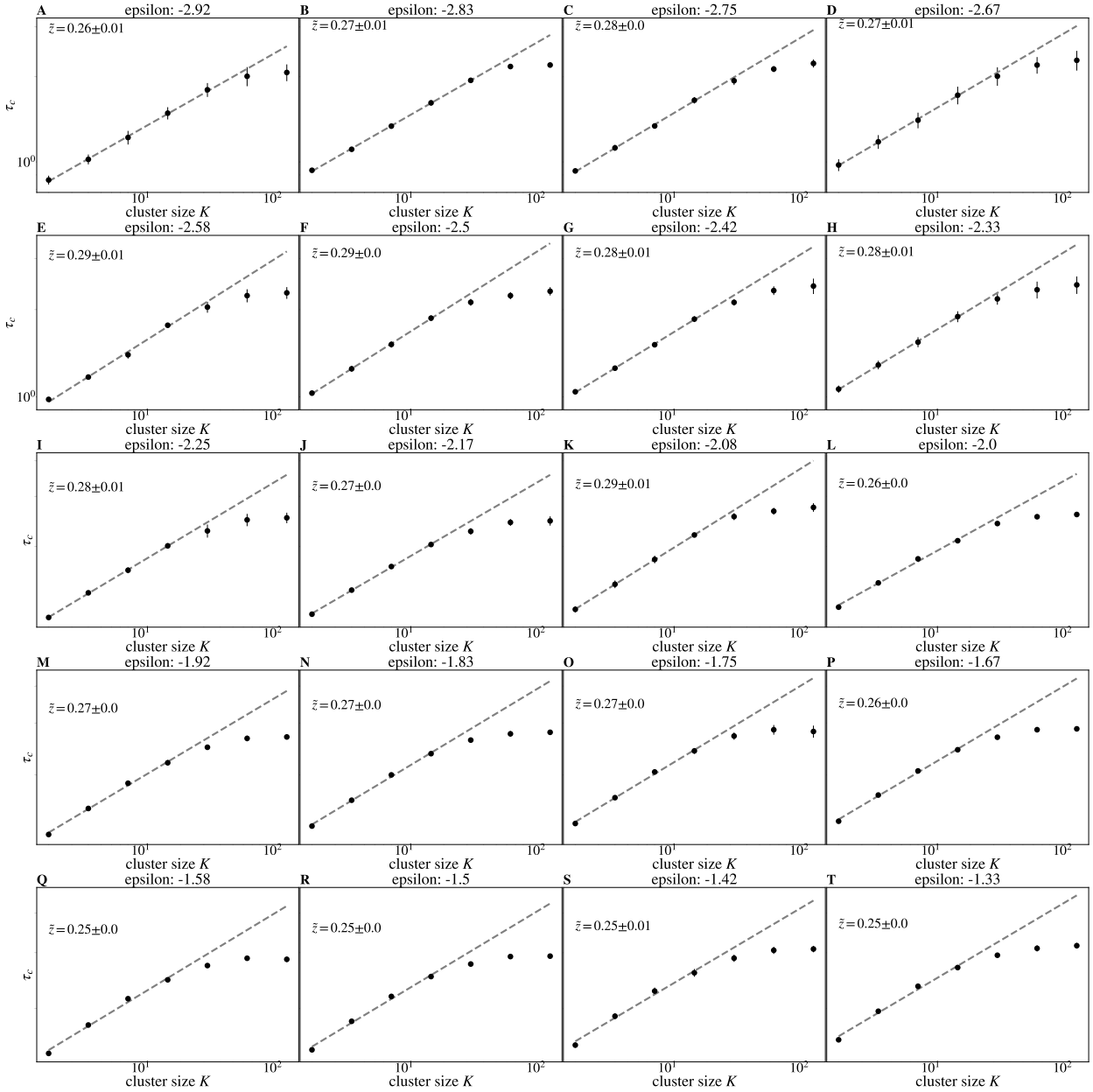

FIG. 71. Time constants τ_c extracted from each curve in FIG 70. Observe behavior obeying $\tau_c \propto K^{\bar{z}}$ for roughly 1 decade. Default simulation parameters with labeled values of ε. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

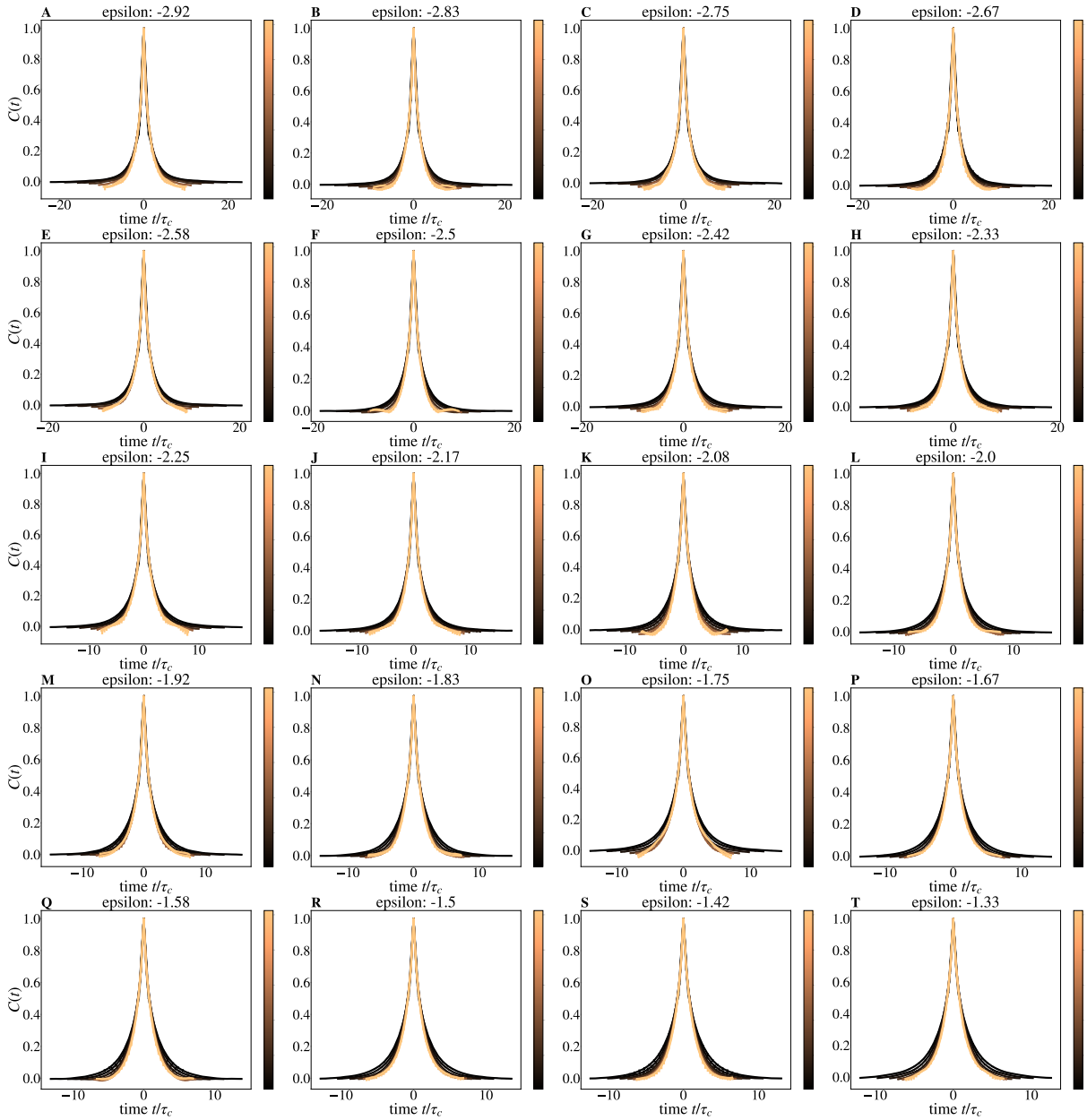

FIG. 72. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$, where time is rescaled by the appropriate τ_c for that coarse-graining iteration. Default simulation parameters with labeled values of ε. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

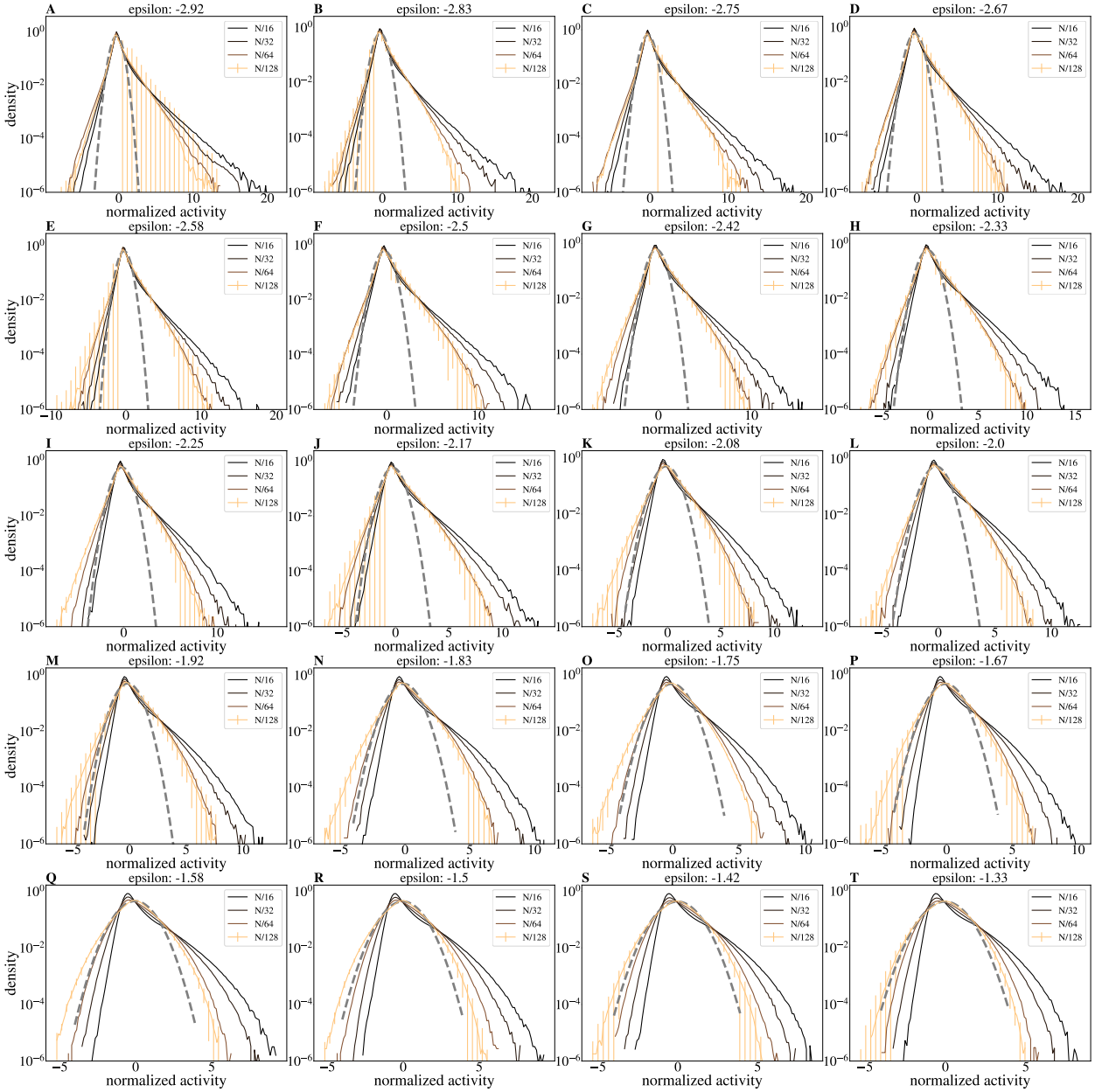

FIG. 73. Distribution of coarse-grained variables for $k = N/16, N/32, N/64, N/128$ modes retained. Default simulation parameters with labeled values of ϵ. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

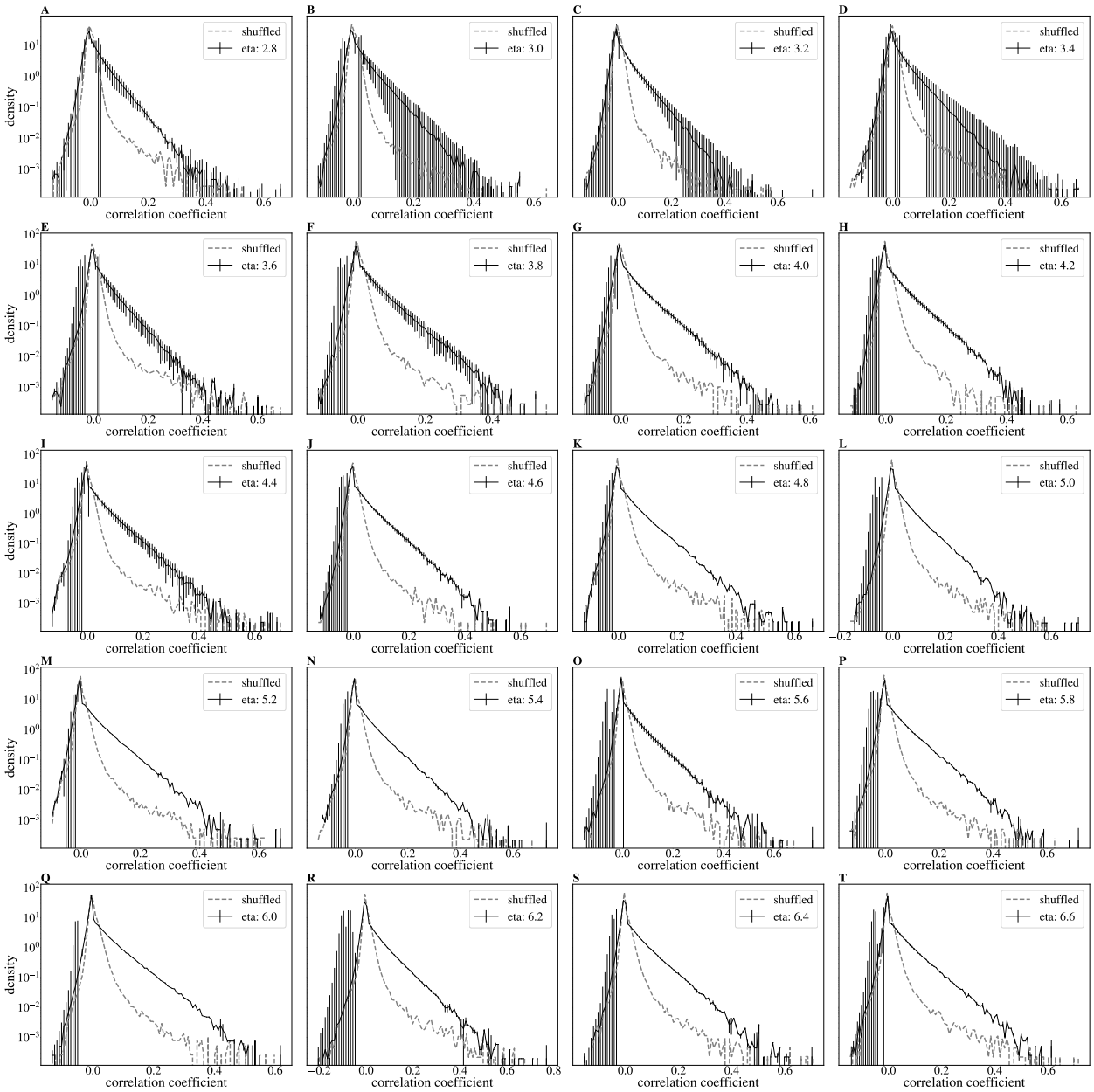

FIG. 74. Distributions of pairwise correlation coefficients. Default simulation parameters with value of η. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

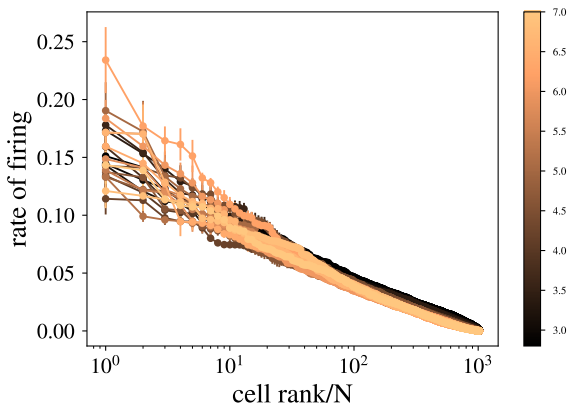

FIG. 75. Rate of firing vs rescaled rank. Colorbar indicates value of η. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

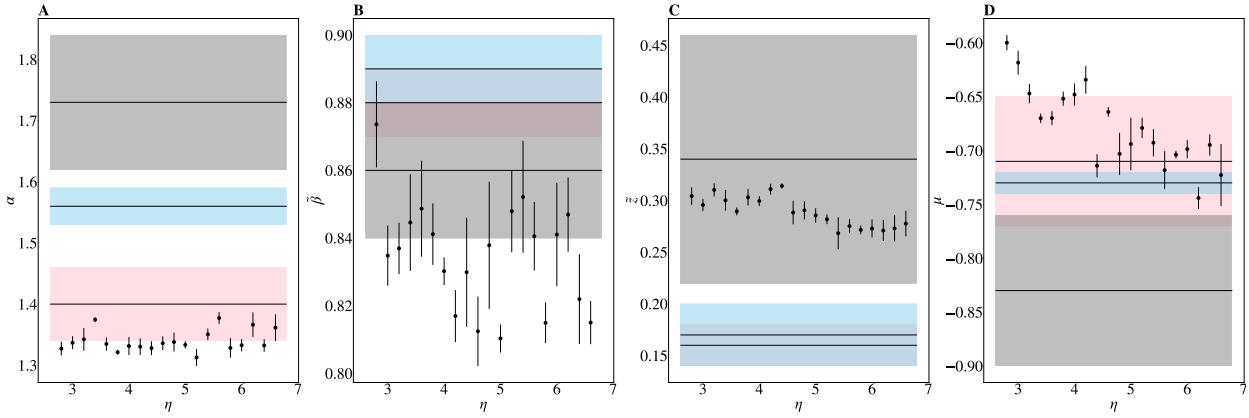

FIG. 76. Each critical exponent, $\alpha, \tilde{\beta}, \tilde{z}, \mu$ vs multiplier η. Results from [1] marked and shaded in gray, pink, and blue. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

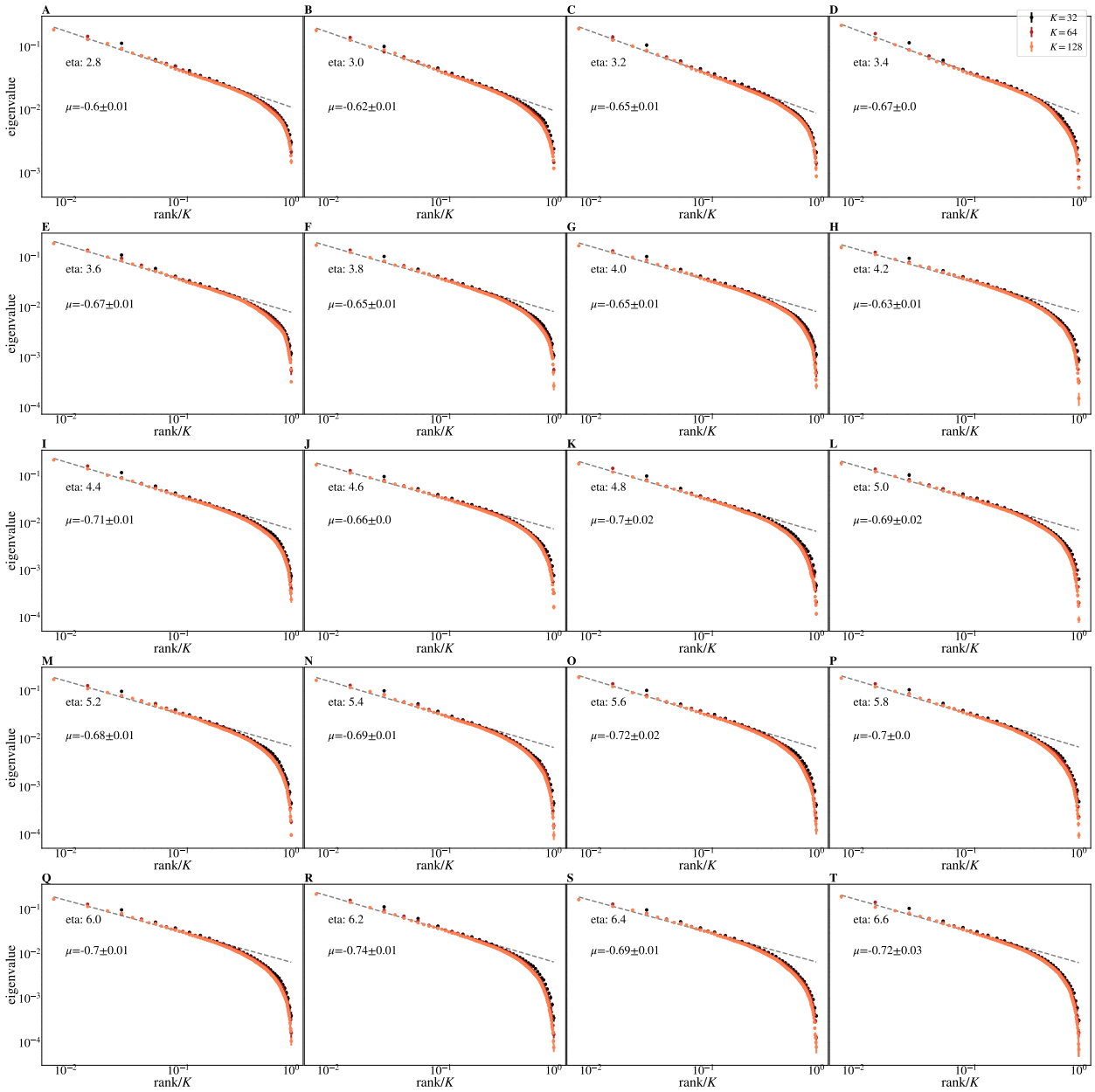

FIG. 77. Average eigenvalue spectrum of Eq 40 for cluster sizes $K = 32, 64, 128$. Default simulation parameters with labeled values of η. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

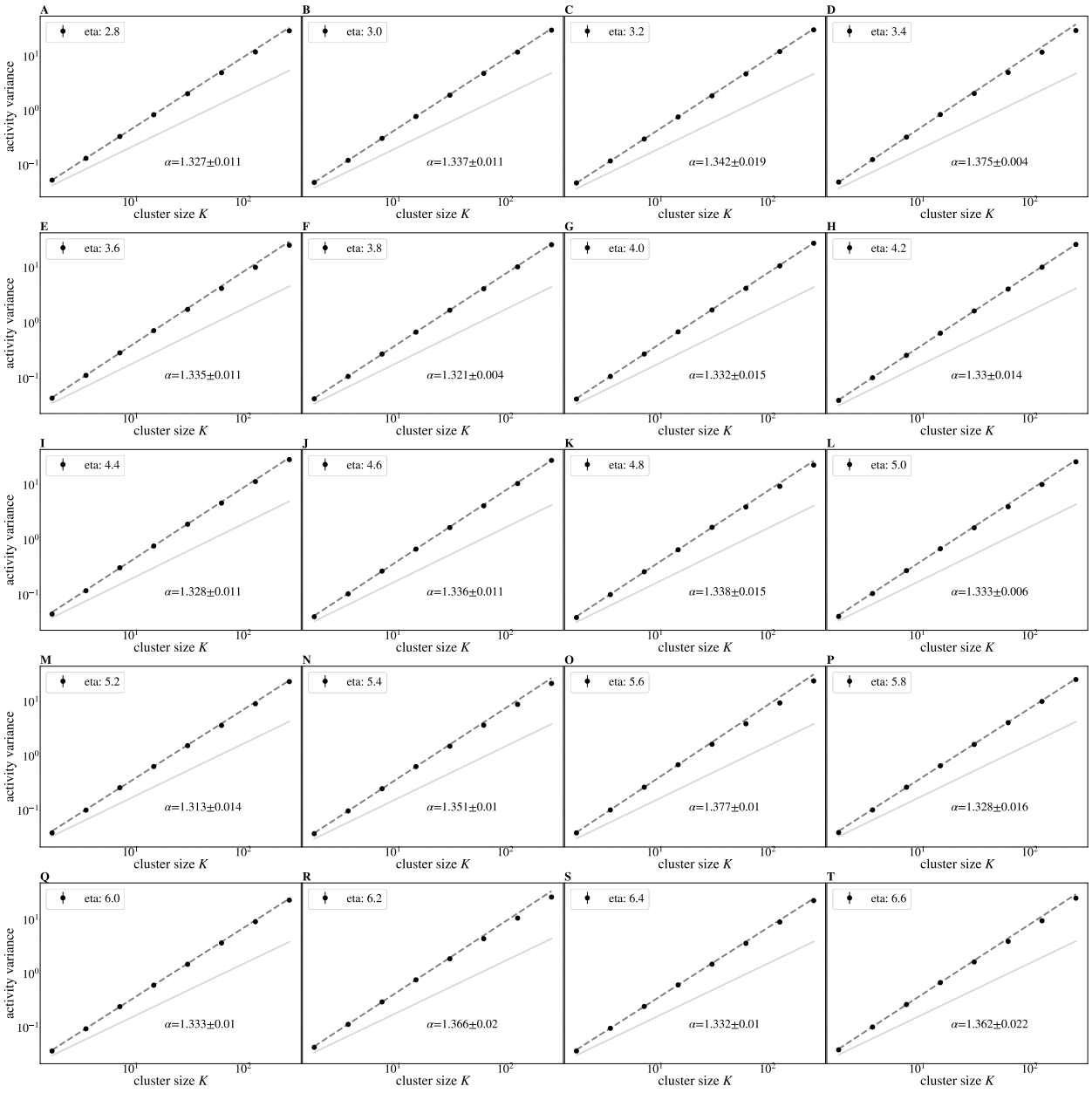

FIG. 78. Activity variance (Eq 33) over coarse-grained variables at each coarse-graining iteration. Default simulation parameters with labeled values of η. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

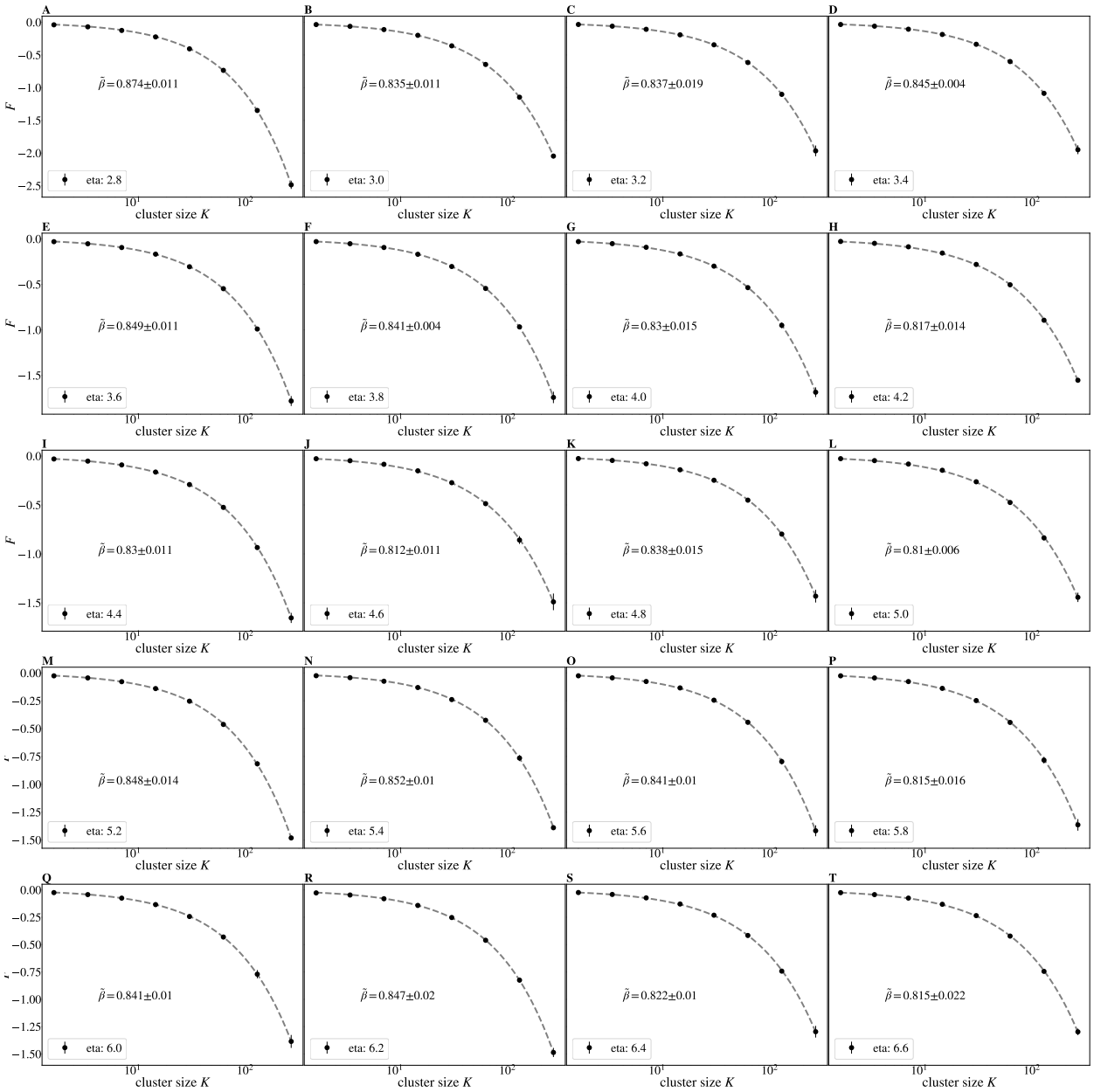

FIG. 79. Average free energy (Eq 34) at each coarse-graining iteration. Default simulation parameters with labeled values of η. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

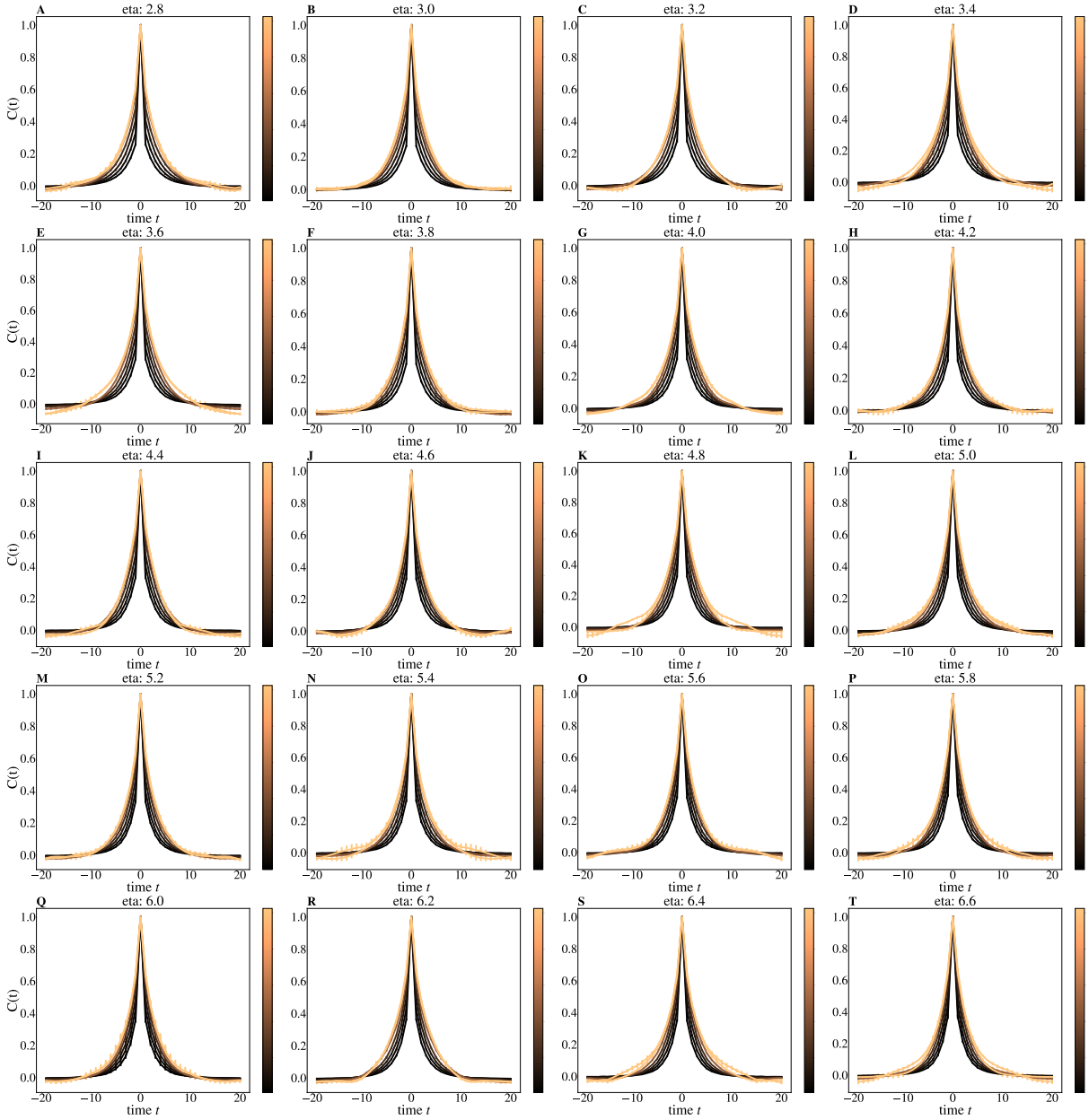

FIG. 80. Average autocorrelation function for cluster sizes $K = 2, 4, ..., 256$ as a function of time, cluster size indicated by colorbar. Default simulation parameters with labeled values of η. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

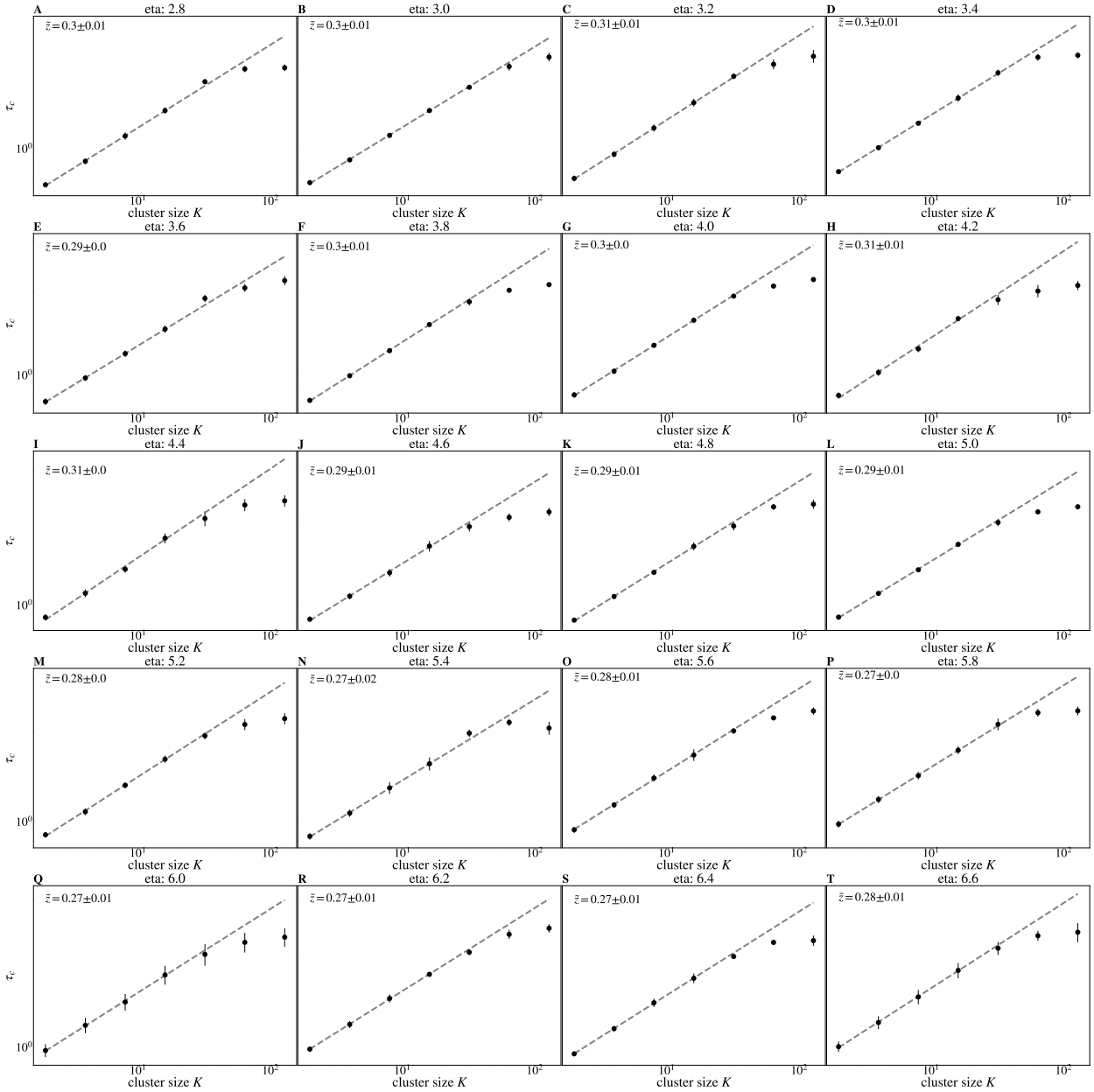

FIG. 81. Time constants τ_c extracted from each curve in in FIG 80, and observe behavior obeying $\tau_c \propto K^{\tilde{z}}$ for roughly 1 decade. Default simulation parameters with labeled values of η. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

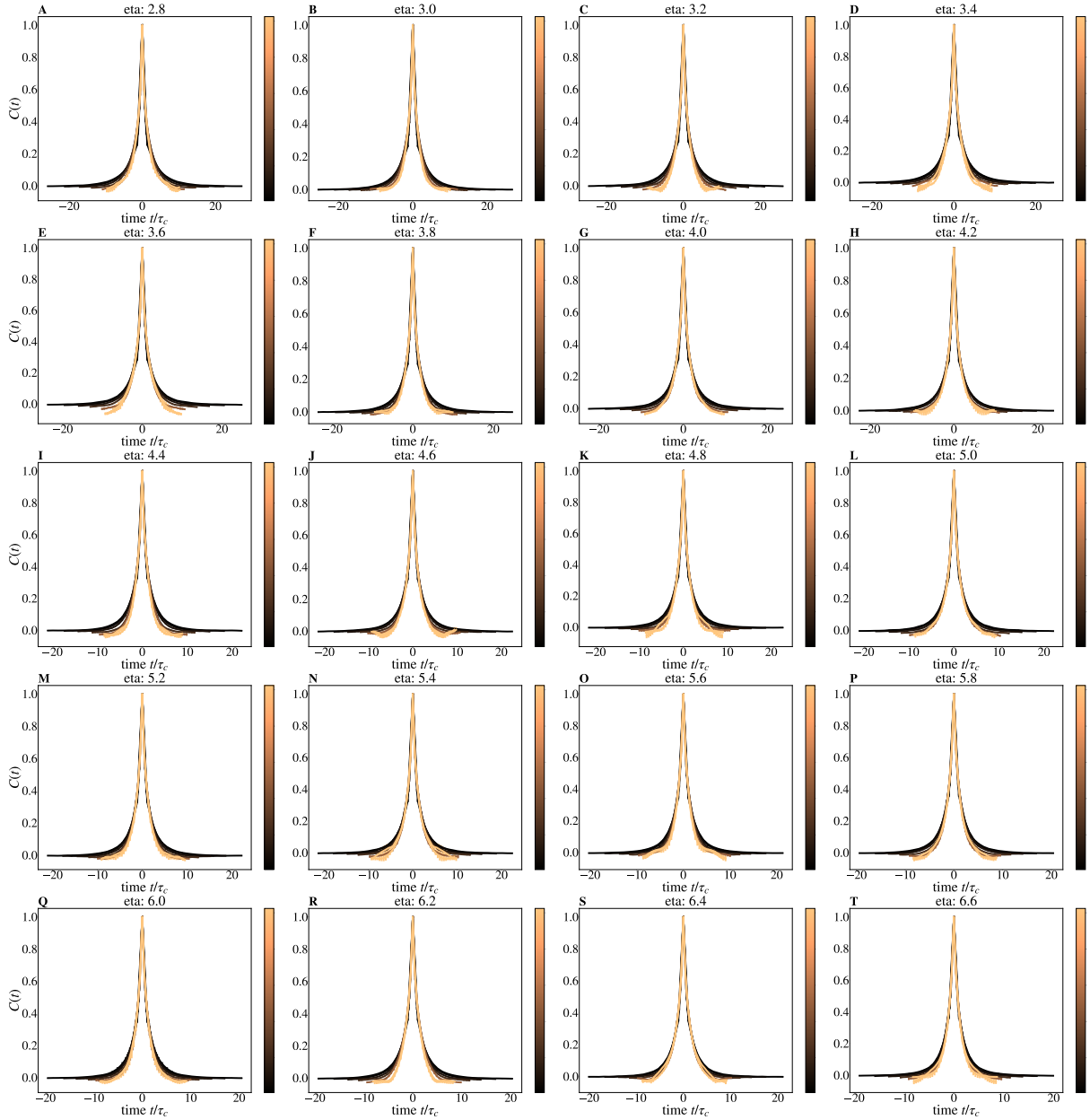

FIG. 82. Average autocorrelation function for cluster sizes $K = 2, 4, \ldots, 256$, where time is rescaled by the appropriate τ_c for that coarse-graining iteration. Default simulation parameters with labeled values of η. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.

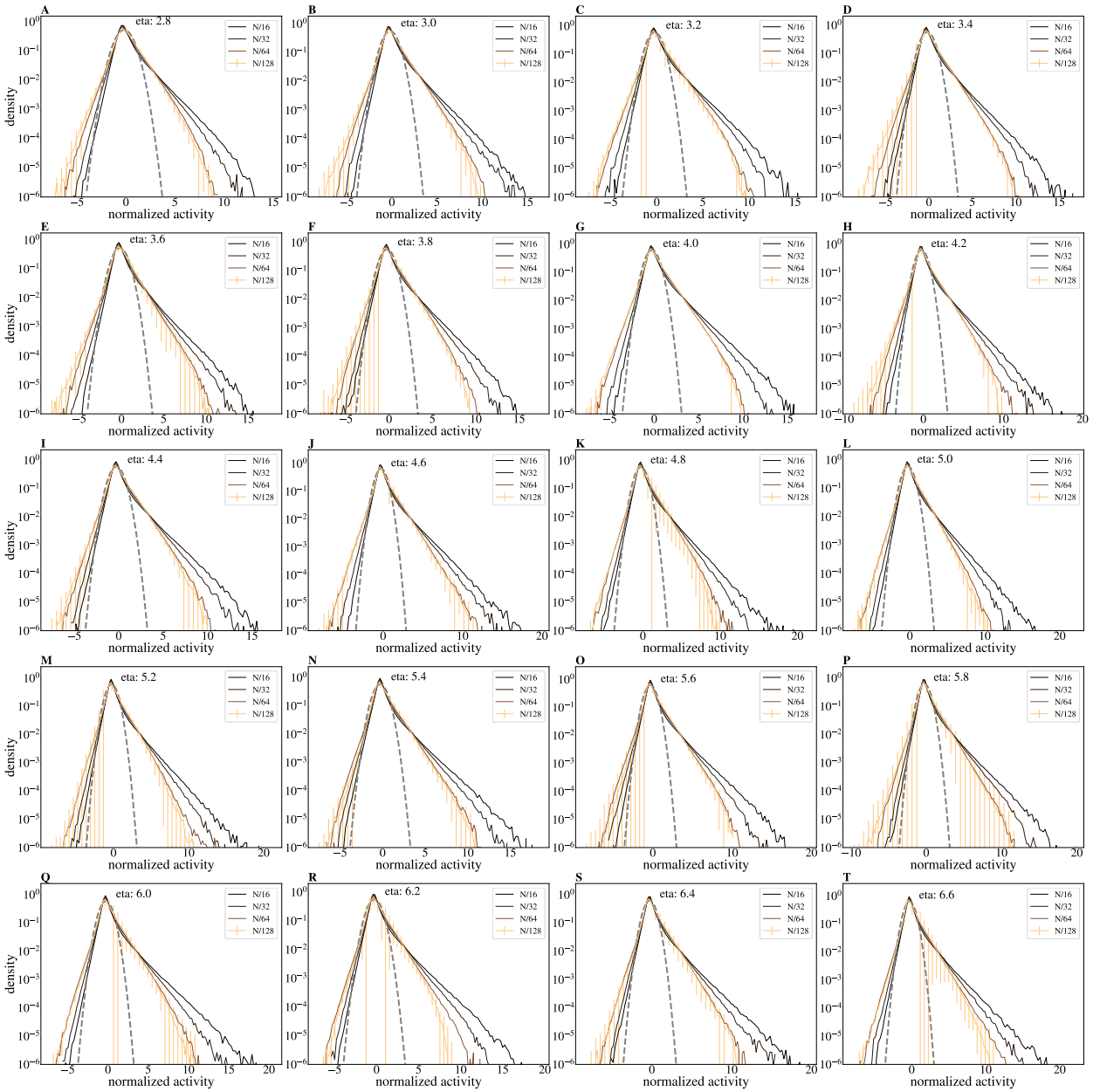

FIG. 83. Distribution of coarse-grained variables for $k = N/16, N/32, N/64, N/128$ retained. Default simulation parameters with labeled values of η. Error bars are standard deviations over randomly selected contiguous quarters of the simulation.