14th International Conference on Structural Safety and Reliability - ICOSSAR'25
1-6 June, 2025 - Los Angeles California, USA
R. Ghanem, E. Johnson, S. Masri and A. Olivier (Eds)

COLLECTIVE BEHAVIOR AND PHASE TRANSITIONS IN REGIONAL SEISMIC RESPONSES

SEBIN OH[1], RAUL RINCON[2], JAMIE E. PADGETT[3], and ZIQI WANG[4]

[1] University of California, Berkeley
Berkeley, CA 94720
sebin.oh@berkeley.edu

[2] Rice University
Houston, TX 77005
raul.rincon@rice.edu

[3] Rice University
Houston, TX 77005
jamie.padgett@rice.edu

[4] University of California, Berkeley
Berkeley, CA 94720
ziqiwang@berkeley.edu

Key words: Regional-scale risk analysis, Regional seismic responses, Collective behavior, Phase transition, Phase diagram

Abstract. Regional seismic risk assessments are typically performed through bottom-up approaches. This approach provides high-resolution responses but often overlooks macroscopic patterns underlying the regional behavior. Moreover, modeling uncertainties and their compounded effects, as well as the interdependencies between hazard and structural performance, further complicate the identification of emergent collective behaviors. We propose a complementary top-down perspective that characterizes regional-scale seismic behavior using concepts from statistical physics. Within a probabilistic simulation framework, we augment conventional models by introducing an additional epistemic uncertainty term to represent model incompleteness. Applying the framework to the residential building portfolio in Milpitas, California, comprising 13,049 buildings, we construct a *phase diagram* that summarizes regional seismic behavior in a hazard–uncertainty space, with earthquake magnitude and epistemic uncertainty as coordinate axes. Two main findings emerge: (1) a *discontinuous phase transition*, representing an abrupt shift from a collective safe state to a collective failure state, occurs as the magnitude exceeds a critical threshold; and (2) a *continuous phase transition*, where the statistical distribution of system-level losses evolves smoothly from a bimodal to a unimodal distribution as epistemic uncertainty grows. Mapping these findings onto a classical statistical physics model clarifies how correlation, heterogeneity, and uncertainty govern systemic fragility: correlations buffer minor hazards but trigger a 'brittle systemic collapse' under larger events; structural diversity enhances 'global ductility'; and excessive epistemic uncertainty blurs the phase boundary, potentially leading to a misrepresentation of regional seismic risk. This study advances a high-level understanding of regional seismic risk by revealing collective behaviors and phase transitions, complementing detailed bottom-up simulations and informing system-level resilience strategies.

1 INTRODUCTION

Understanding and mitigating the impact of earthquakes on civil infrastructure systems is an urgent challenge that demands systematic and scalable approaches. Current regional-scale seismic risk assessment methods primarily rely on bottom-up simulation models, such as scenario- and component-based approaches, which predict the responses of individual components under specific hazard scenarios and aggregate these results to estimate system-level performance [1, 2]. While these methods offer high-resolution predictions, they are computationally intensive, limiting their scalability, particularly given that regional analyses often involve millions of components. Moreover, bottom-up approaches make it difficult to extract macroscopic patterns across numerous simulations, rendering them insufficient for intelligible regional risk analysis and for developing comprehensive seismic risk mitigation strategies.

As a complementary paradigm, a top-down approach seeks to understand regional seismic responses through governing principles and universal macroscopic patterns. For example, if structural damage states are independent across a region, the Central Limit Theorem (CLT) applies, indicating that a global risk metric based on the sum of individual damage states would converge to a Gaussian distribution without the need for numerous simulations. However, when correlations between damage states are considered, the CLT no longer holds, and collective behaviors deviating from Gaussian distributions can emerge. With the growing emphasis on regional-scale risk assessment for earthquakes, systematically characterizing these collective responses across different earthquake scenarios becomes essential, offering a high-level perspective that complements traditional component- and scenario-based simulations.

The emergence of collective behavior is a traditional province of statistical physics [3]. In that discipline, universal patterns that emerge regardless of the microscopic details of a system are of primary interest, including phase transitions of matter. The complex nature of regional seismic responses, arising from both geological complexity and structural randomness, suggests the potential applicability of a statistical physics perspective to regional seismic risk analysis. In fact, Oh et al. (2024) recently showed that the Ising model, a seminal model in statistical physics, can be adapted to describe regional seismic responses by treating the damage state of each structure as a spin and the seismic hazard as an external magnetic field [4].

To capture such collective behavior, regional-scale correlations have been proposed between seismic intensities [5] and among structures of expected similar performance [1]. Beyond correlations of the physical phenomena, effects from modeling uncertainties have been shown to also induce correlated modifications in performance estimates. For example, Rincon and Padgett (2025) showed that varying levels of knowledge about the structural fragility of exposed systems (due to modeling uncertainty or incompleteness) induce a compounded effect over the regional-scale outcomes [6].

Building on these insights, this study investigates collective behavior and phase transitions in regional seismic responses through the lens of statistical physics. Using a probabilistic simulation framework applied to a residential building portfolio in Milpitas, California, we systematically vary earthquake magnitude and model uncertainty to uncover emergent macroscopic patterns. The resulting diagram, which summarizes regional seismic responses across these two variables, reveals both discontinuous and continuous phase transitions, closely resembling the behavior of classical statistical physics models, such as the Ising model. These findings offer a new perspective on regional seismic risk and underscore the potential of statistical physics methods for understanding and managing systemic vulnerabilities in civil infrastructure systems.

2 REGIONAL-SCALE SEISMIC SIMULATION

This study adopts the concept of a *safety margin* $\boldsymbol{X}$ to describe seismic responses, defined as $\boldsymbol{X} = \boldsymbol{C} - \boldsymbol{D}$, where $\boldsymbol{C}$ and $\boldsymbol{D}$ are the capacity and demand vectors, respectively. Each element of $\boldsymbol{X}$ quantifies how far an individual structure is from failure, with negative values denoting failure and positive values denoting survival. In practice, the safety margin is often expressed in logarithmic form, $\boldsymbol{X} = \ln \boldsymbol{C} - \ln \boldsymbol{D}$, where $\boldsymbol{C}$ and $\boldsymbol{D}$ are functions of either intensity measures (IMs) or engineering demand parameters (EDPs). This section first presents the modeling approaches for seismic demand and structural capacity, and then introduces the overall framework built upon these models.

2.1 Demand modeling

Seismic demand modeling can be performed using various methods, including ground motion prediction equations (GMPEs), which estimate IMs at different sites for a given earthquake scenario [7]. Structural EDPs can also serve as demand measures by combining GMPEs with IM–EDP regression models that predict EDPs from given IMs [8, 9]. Nonlinear time history analysis of structural models subjected to ground motions provides another approach for estimating accurate EDPs. Advances in computational resources have also enabled physics-based ground motion simulations incorporating wave propagation through finite element analysis [10, 11]. However, applying such simulations at a regional scale, involving tens of thousands of structures, remains computationally prohibitive for individual researchers. Consequently, this study adopts the conventional GMPE-based approach to estimate seismic demand across structural locations, using peak ground acceleration (PGA) as the demand measure.

The ground motion IM at site i under earthquake scenario j is typically modeled as

$$\ln\left(IM_{ij}\right) = \ln\left(\overline{IM}_{ij}\right) + \sigma\varepsilon_{ij} + \tau\eta_j \tag{1}$$

where IM_{ij} denotes the IM at site i due to earthquake j, and $\overline{IM}_{ij}$ is the median IM predicted by a GMPE, which depends on earthquake characteristics such as magnitude, distance, spectral period, and local site conditions. The terms ε_{ij} and η_j represent the normalized intra-event and inter-event residuals, respectively, and are both assumed to follow standard normal distributions with zero mean and unit standard deviation. The corresponding standard deviation terms, σ and τ, quantify the variability within and between events, respectively. These coefficients typically depend on the spectral period and may also vary with earthquake magnitude and the distance from the rupture to the site. The term $\sigma\varepsilon_{ij}$ is known as the intra-event residual, while $\tau\eta_j$ represents the inter-event residual. As indicated by the absence of the site index i, the inter-event residual η_j is constant across all sites subjected to the same earthquake scenario. In addition, to account for the correlation between the IMs across different sites, the correlation between the intra-event residuals, $\rho_{\varepsilon_{i_1 j}\varepsilon_{i_2 j}}$, is calculated as

$$\rho_{\varepsilon_{i_1 j}\varepsilon_{i_2 j}}(h) = \exp\left(-3h/R\right) \tag{2}$$

as proposed by Jayaram and Baker (2009) [5], where h denotes the distance in kilometers between sites i_1 and i_2, and R is a parameter that controls the rate at which spatial correlation decays with distance, for which a value of $R = 8.5$ is adopted in this study based on PGA.

The GMPE proposed by Atkinson and Boore (1995) [12] is adopted in this study to estimate the median IM values, $\overline{IM}_{ij}$, at different sites. The model takes the following form:

$$\log\left(\overline{IM}_{ij}\right) = c_1 + c_2\left(M_w - 6\right) + c_3\left(M_w - 6\right)^2 - \log R - c_4 R \tag{3}$$

where M_w is the earthquake moment magnitude, R denotes the hypocentral distance, and c_1 through c_4 are empirically determined regression coefficients. The model also provides recommended values for the inter-event and intra-event standard deviations, τ and σ, which depend on the specific IM of interest. In this study, the values $c_1 = 3.79$, $c_2 = 0.298$, $c_3 = -0.0536$, $c_4 = 0.00135$, $\tau = 0.18$, and $\sigma = 0.20$ are adopted, corresponding to PGA. The soil amplification factor of 2 is also adopted to account for site effects associated with the soil conditions of the target region, which is detailed in Section 3.

2.2 Capacity modeling

For structural capacity, the nominal values specified in design codes or guidelines are typically used. These values are derived through a combination of empirical data, analytical modeling, and expert judgment [13, 14, 15]. In many cases, such codes and guidelines also provide fragility curves for different structural types, which effectively represent the distribution of structural capacity around the nominal value. These curves estimate the probability of failure, or the exceeding probability for a specified damage state, for a given IM or EDP. Specifically, a fragility curve is commonly modeled using a lognormal distribution as follows:

$$\mathbb{P}\left(DS \geq ds_i \mid D = d\right) = \Phi\left(\frac{\ln d - \ln \theta_i}{\beta_i}\right) \tag{4}$$

where ds_i denotes the i-th damage state, D is the seismic demand at the location of the structure, $\Phi(\cdot)$ is the standard normal cumulative distribution function, θ_i is the median demand causing damage state ds_i or greater, and β_i is the logarithmic standard deviation representing the uncertainty in that threshold. Note that the fragility expression in Eq. (4) can be equivalently interpreted in terms of structural capacity as [16]

$$\mathbb{P}\left(DS \geq ds_i \mid D = d\right) = \mathbb{P}\left(C_i < D\right) \tag{5}$$

where C_i represents the structural capacity associated with the i-th damage state and is assumed to follow a lognormal distribution with median θ_i and logarithmic standard deviation β_i, i.e., $C_i \sim \mathcal{LN}(\theta_i, \beta_i)$. In case when only the binary states of a structure, i.e., failure and non-failure, are of interest rather than detailed performance states, we may exclude the subscript i, leading to

$$C \sim \mathcal{LN}(\theta, \beta)\,, \tag{6}$$

where the definition of failure depends on a particular damage state of interest.

While fragility curves for individual structures have been extensively developed over decades of research, relatively fewer studies have explored the extension of fragility curve-based approaches to groups of structures or investigated the correlation of structural capacities [1, 17]. Building on the framework proposed by Baker et al. (2024), which decomposes variability in structural component-level capacity into multiple sources and models correlation arising from common source effects by assigning weights to each source [18], this study proposes to decompose the structure-level capacity as follows:

$$\ln C = \ln \theta + \beta\left(\sqrt{\boldsymbol{w}_{\text{dep}}} \cdot \boldsymbol{\varepsilon}_{\text{dep}} + \sqrt{w_{\text{ind}}}\varepsilon_{\text{ind}}\right) \tag{7}$$

where $\boldsymbol{w}_{\text{dep}}$ is a weight vector associated with sources of correlation, $\sqrt{\boldsymbol{w}_{\text{dep}}}$ denotes the vector whose components are the element-wise square roots of the entries in $\boldsymbol{w}_{\text{dep}}$, and $\boldsymbol{\varepsilon}_{\text{dep}}$ is a multivariate standard normal random vector. Their dot product, $\sqrt{\boldsymbol{w}_{\text{dep}}} \cdot \boldsymbol{\varepsilon}_{\text{dep}}$, captures the correlated component of

structural capacities. The second term inside parentheses, $\sqrt{w}_{\text{ind}}\varepsilon_{\text{ind}}$, represents the structure-specific variability, quantifying how independent a structural capacity is relative to others. Here, w_{ind} is the weight assigned to the independent source, and ε_{ind} is a standard normal random variable. Note that the sum of the elements in $\boldsymbol{w}$dep and $w$ind should equal one to ensure that the total standard deviation of $\ln C$ remains β, as specified in Eq. (6).

For the sake of illustrating its use, we present the application of the correlation model for building portfolios. Considering the accessible building inventory information, this study adopts the following sources of uncertainty: design code level w_{code}, structural properties (including structural material and structural type) w_{str}, occupancy class w_{occ}, and number of stories w_{num}. The corresponding weights are defined as

$$\boldsymbol{w}_{\text{dep}} = [w_{\text{code}}, w_{\text{str}}, w_{\text{occ}}, w_{\text{num}}, w_{\text{ind}}]^{\text{T}} = [0.15, 0.3, 0.05, 0.35, 0.15]^{\text{T}}. \tag{8}$$

During capacities' sampling, each simulation shares the same $\boldsymbol{\varepsilon}_{\text{dep}}$ values across all structures influenced by the same uncertainty sources; additionally, independent values ε_{ind} are sampled individually for each structure. For example, if two structures share the same design code level and structural properties but differ in occupancy class and number of stories, they will share the same $\varepsilon_{\text{code}}$ and ε_{str} values in each simulation. This approach introduces correlation among the capacities by linking them through the shared components.

2.3 Framework formulation considering uncertainty

As described in Sections 2.1 to 2.2, this study adopts a GMPE-based approach to estimate seismic demand and a fragility curve-based approach to simulate structural capacity. Since both approaches assume that the logarithmic values of demand and capacity follow Gaussian distributions, the safety margin, defined as $\boldsymbol{X} = \ln \boldsymbol{C} - \ln \boldsymbol{D}$, also follows a Gaussian distribution:

$$\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu_X}, \boldsymbol{\Sigma_X}) \tag{9}$$

where $\boldsymbol{\mu_X}$ and $\boldsymbol{\Sigma_X}$ denote the mean vector and covariance matrix of $\boldsymbol{X}$, respectively. These can be expressed in terms of the statistics of the logarithmic capacity $\ln \boldsymbol{C}$ and demand $\ln \boldsymbol{D}$ as

$$\boldsymbol{\mu_X} = \boldsymbol{\mu_C} - \boldsymbol{\mu_D}, \quad \boldsymbol{\Sigma_X} = \boldsymbol{\Sigma_C} + \boldsymbol{\Sigma_D}, \tag{10}$$

assuming that $\boldsymbol{C}$ and $\boldsymbol{D}$ are independent. Here, $\boldsymbol{\mu_C}$ and $\boldsymbol{\Sigma_C}$ represent the mean vector and the covariance matrix of the logarithmic capacity, respectively, and are obtained from Eq. (7), while $\boldsymbol{\mu_D}$ and $\boldsymbol{\Sigma_D}$ denote the corresponding statistics of logarithmic demand, computed using Eq. (1). While the Gaussian distribution in Eq. (9) arises directly from the normality assumptions made for capacity and demand modeling approaches adopted, the assumption remains reasonable in most practical applications. This is because the Gaussian distribution is the least biased distribution when prior knowledge is limited to the second-order central moments, i.e., the mean vector and the covariance matrix [19].

Although the realism of the overall framework represented by Eq. (9) can be improved by conducting a detailed examination of both capacity and demand models and refining their associated statistics, it still possesses inherent incompleteness. For instance, when fragility curves derived from HAZUS-MH 5.1 are used, the statistical properties embedded in those curves are limited to the structural characteristics (e.g., material, structural type, design code level, number of stories) considered during their development, which may not fully represent all factors influencing the behavior of real structures. Similarly, when seismic demand is estimated using a GMPE, the prediction is based

on the seismological parameters incorporated in the GMPE (e.g., magnitude, hypocentral distance, soil type), which often fail to capture the full complexity of ground motion. Moreover, these two submodels are integrated for regional analysis, creating complex interactions of the modeling uncertainties [6].

To account for the inherent incompleteness of the modeling framework, an additional uncertainty term is introduced into the safety margin as follows:

$$\boldsymbol{X} \sim \mathcal{N}\left(\boldsymbol{\mu_X},\, \boldsymbol{\Sigma_X} + \sigma_0^2 \boldsymbol{I}\right), \tag{11}$$

where σ_0 represents the level of uncertainty arising from the incompleteness of the modeling framework, and $\boldsymbol{I}$ is the identity matrix, reflecting the lack of information about potential correlations in this uncertainty across different components. This additive term collectively captures the limitations of both the demand and capacity models. Incorporating broader structural and seismological characteristics may help reduce σ_0 by improving the completeness of the overall model. Notably, the case of $\sigma_0 = 0$ implies that the framework fully accounts for all relevant aspects of the problem, thereby enabling the recovery of the ground truth through the reduction of epistemic uncertainty within it. The regional-scale seismic simulation can then be performed by generating the safety margin vector and assigning binary damage states based on its values: failure if $X_i < 0$ and non-failure if $X_i > 0$, where i denotes the index of structures. Section 3 shows how this framework is applied to simulate regional-scale seismic responses and to discover collective behavior emerging from them.

3 COLLECTIVE BEHAVIOR OBSERVED IN REGIONAL SEISMIC RESPONSES

This section applies the framework illustrated in Section 2 to Milpitas, California. This city is predominantly a residential area in the southern part of the San Francisco Bay Area (Figure 1). A total of 13,049 residential buildings are selected from 14,280 buildings in the city to demonstrate how the residential buildings respond to earthquakes in a collective sense.

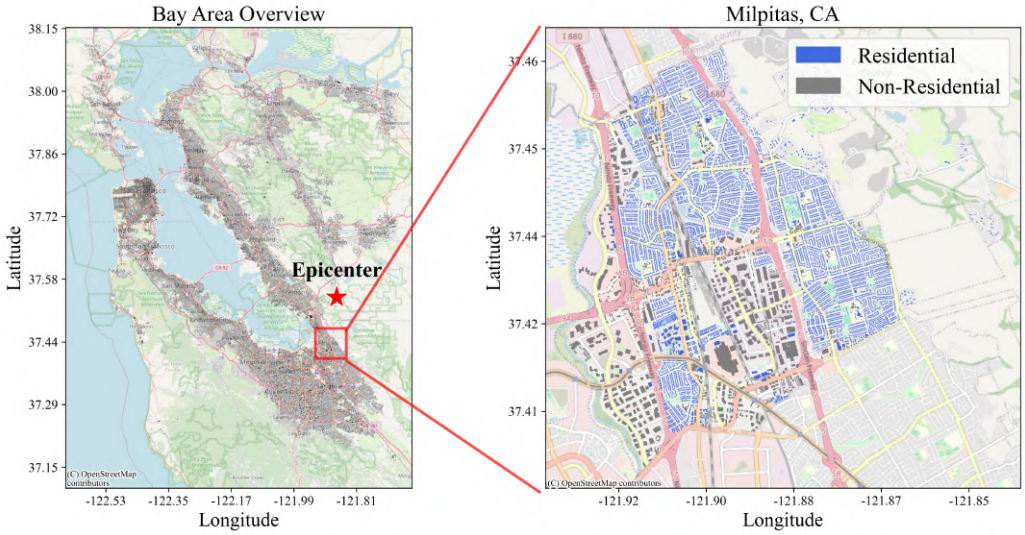


Figure 1: Target region: Milpitas, California. A total of 13,049 residential buildings are selected from 14,280 buildings in the city to demonstrate collective behavior in a residential area.

The building inventory data is retrieved from the DesignSafe Data Depot [20]. This dataset includes information such as geographical location, number of stories, year built, occupancy class,

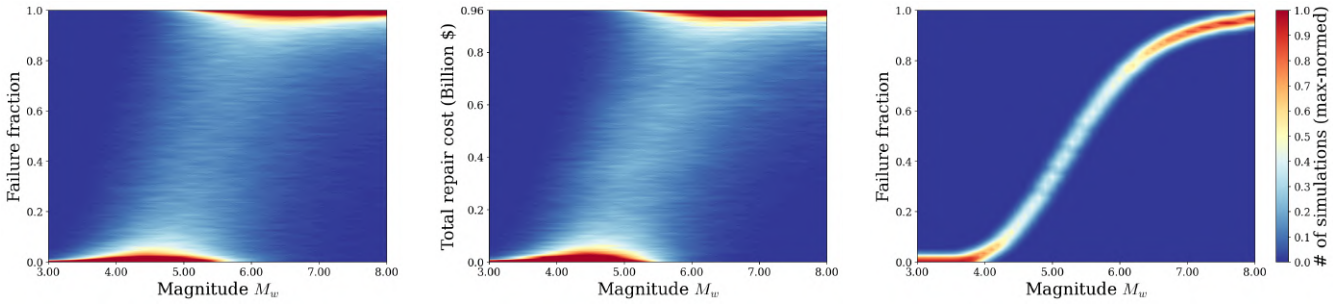
structural type, and repair cost. A PGA-based fragility curve for the moderate damage state is used, with parameters adopted from HAZUS-MH 5.1 [13], which categorizes buildings based on structural type, number of stories (low-, mid-, and high-rise), and year built (pre-, low-, moderate-, and high-code). The GMPE parameters are adopted from the literature [5, 12], as described in Section 2.1. The earthquake is assumed to originate from a point source with an epicenter located on the Hayward Fault (see the left plot in Figure 1) at a depth of 10 km.

To quantify regional risk for each simulation, two global metrics are used: the failure fraction and the total repair cost. The failure fraction is defined as the proportion of buildings reaching or exceeding the moderate damage state, while the total repair cost is calculated as the sum of the repair costs of the failed buildings, assuming that the repair costs retrieved from the DesignSafe Data Depot [20] correspond to the moderate damage state. The following sections explore two distinct types of collective behavior by varying the earthquake magnitude M_w and the uncertainty level of the analysis framework σ_0. For each (M_w, σ_0) analysis scenario, a total of 10,000 simulations are conducted.

3.1 Collective behavior across earthquake magnitudes

To investigate how the residential building portfolio in the target region behaves under different earthquake magnitudes, simulations are conducted across 51 different M_w values, ranging from 3.0 to 8.0 in increments of 0.1. No additional uncertainty is considered in this demonstration (i.e., $\sigma_0 = 0$), under the assumption that the current modeling framework sufficiently accounts for the relevant factors governing regional seismic responses.

Figure 2 illustrates the collective behavior exhibited by the target region under varying earthquake magnitudes. The color indicates the maximum-normalized number of simulations. Each plot can be effectively interpreted as an aggregation of histograms at different magnitude values. Notably, two isolated red islands in the lower left and upper right of both plots, Figures 2a and 2b, signify an abrupt transition from a *collective safe state* to a *collective failure state*. This abrupt transition is due to the correlation between the damage states of buildings, as confirmed by comparison with the independent case shown in Figure 2c. Figure 2c is obtained following the same procedures as for Figure 2a, but under the independence assumption between building damage states (i.e., $\boldsymbol{\Sigma}_{\boldsymbol{X}}$ is a diagonal matrix), where the transition from a collective safe state to a collective failure state occurs smoothly without discontinuity.



(a) (b) (c)

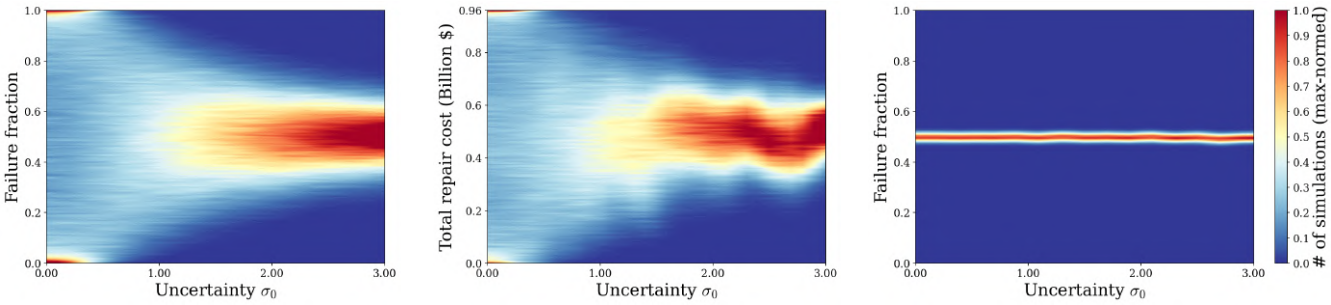
Figure 2: Collective behavior exhibited in (a) the failure fraction and (b) the total repair cost under varying earthquake magnitudes. The case of independent damage states is shown in (c) for comparison.

3.2 Collective behavior across uncertainty levels

This section explores the collective behavior associated with the uncertainty inherent in the analysis framework by varying σ_0 from 0.0 to 5.0 in increments of 0.1, while keeping M_w fixed. Although any earthquake magnitude could be used, $M_w = 5.36$ is chosen as it marks the transition from a collective safe state to a collective failure state (see Figure 2). All other experimental settings are identical to those described in Section 3.1.

The outcomes under varying uncertainty levels are illustrated in Figure 3. To improve visualization, only the σ_0 range of $[0.0, 3.0]$ is shown, extracted from the full range of $[0.0, 5.0]$. Notably, in Figure 3a, the failure fraction exhibits two peaks at the extremes (0.0 and 1.0) at low uncertainty, indicating a bimodal distribution under low uncertainty levels, which then smoothly transitions to a unimodal distribution as uncertainty increases. A similar trend is observed in the total repair cost (Figure 3b), although it appears weaker due to additional incoherency across building repair costs. The polarization of the failure fraction at low uncertainty is attributed to the correlation between the damage states, which causes the portfolio to be collectively safe or failed. At a certain magnitude (around $M_w = 5.36$ in this example), seismic demands and capacities balance, thus the probabilities of a collective safe state and a collective failure state become comparable.

The emergence of a unimodal distribution can be explained by the Central Limit Theorem (CLT); as uncertainty increases, the safety margins (and thus the corresponding damage states) become increasingly independent (see Eq. (9)), and the CLT applies to the failure fraction that essentially represents the normalized sum of the damage states. Indeed, when damage states are assumed to be independent, as shown in Figure 3c, the CLT holds across all uncertainty levels, and the distribution remains effectively constant as a Gaussian distribution centered at a failure fraction equals 0.5.



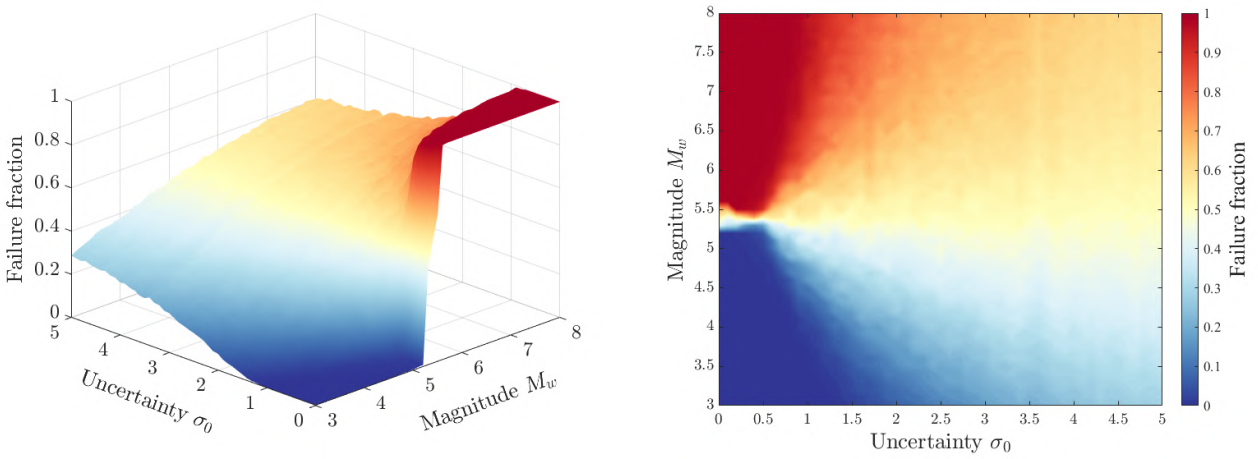
(a) (b) (c)

Figure 3: Collective behavior exhibited in (a) the failure fraction and (b) the total repair cost under varying uncertainty. The case of independent damage states is shown in (c) for comparison.

3.3 Phase diagram for the target region

Building on the collective behaviors described in the previous sections, this section introduces a *phase diagram* for the target region under seismic hazards. For each (M_w, σ_0) pair, where $M_w \in [3.0, 8.0]$ and $\sigma_0 \in [0.0, 5.0]$, 10,000 simulations are conducted, resulting in a total of $51 \times 51 \times 10{,}000 = 26{,}010{,}000$ simulations. To enhance visualization, only the peak values from the heat maps (e.g., Figures 2a and 3a) are extracted and stacked to construct the phase diagram. In other words, each point in the diagram represents the mode (i.e., the most probable value) of the 10,000 simulations under a given (M_w, σ_0) scenario.

The phase diagram is shown in three dimensions in Figure 4a and summarized in two dimensions in Figure 4b. Note that the cross-sections of Figure 4a along the planes $\sigma_0 = 0$ and $M_w = 5.36$ effectively represent Figures 2a and 3a, respectively. In the phase diagrams, redder areas denote scenarios closer to a collective failure state, while bluer areas correspond to collective safe states. The white area separating the red and blue areas along $\sigma_0 \in [0, 0.5]$ in Figure 4b represents the phase boundary between the two states.



(a) (b)

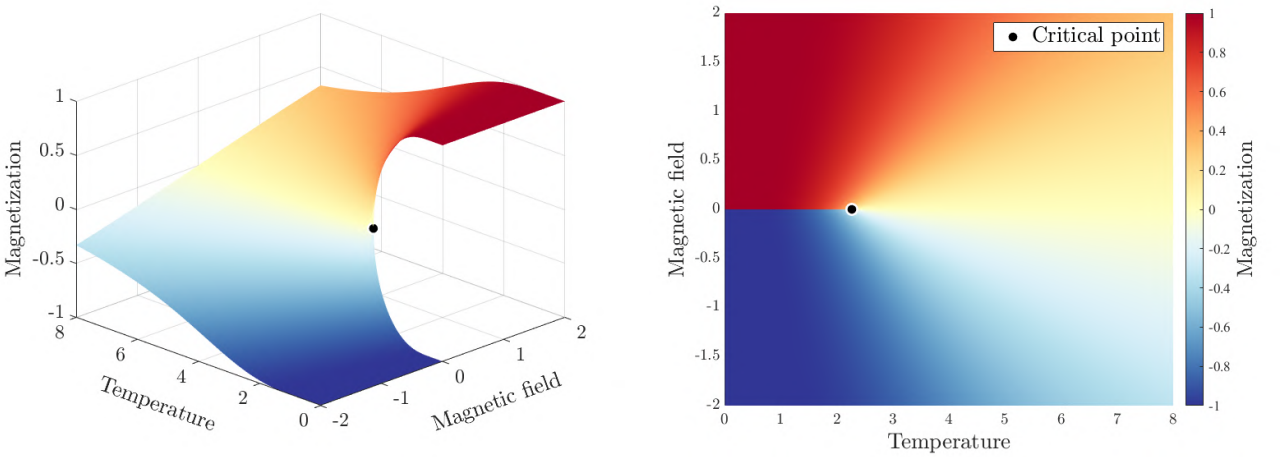
Figure 4: Phase diagram illustrating the seismic response of residential buildings in Milpitas, California, shown in (a) three dimensions and (b) two dimensions.

This phase diagram suggests valuable insights into regional-scale seismic responses. First, correlation makes the system brittle; while it buffers against small hazards, it can trigger a system-wide failure under large hazards. Second, improving the reliability of individual buildings does not necessarily enhance overall system ductility. Instead, increasing heterogeneity among structures (e.g., diverse structural types, occupancy classes, and construction practices) can mitigate the effects of correlated damage states and support more resilient community design. Finally, a narrow modeling framework that incorporates only limited information on hazard and structural performance may lead to significant misrepresentation of seismic risk. It is therefore important to maintain uncertainty below a certain threshold, beyond which the statistical properties of global resilience metrics may change fundamentally. In particular, the transition magnitude separating the collective safe and failure states is critical for seismic risk evaluation, but large uncertainty can obscure its identification.

4 STATISTICAL PHYSICS INTERPRETATION OF THE COLLECTIVE BEHAVIOR

This section interprets the collective behaviors observed in Section 3 through the lens of statistical physics. While the Ising model, a seminal framework for studying critical phenomena and phase transitions, is adopted here, the insights are not limited to this specific model. The Ising model was originally conceived to explain ferromagnetism using binary spins (up and down) [3], and its phase diagram is shown in Figure 5. At low temperatures, magnetization, a global metric of the model, increases abruptly as the external magnetic field crosses zero, while at high temperatures, it increases gradually. This behavior is known as a *discontinuous phase transition*, reflecting the discontinuity

at the transition point. More interestingly, when the external magnetic field is zero, magnetization remains zero at high temperatures but spontaneously shifts to either $+1$ or -1 as the temperature drops below a critical value. This phenomenon, called *spontaneous symmetry breaking*, indicates that the system's inherent symmetry under zero magnetic field is broken due to thermal fluctuations. The associated transition is a *continuous phase transition*, where no abrupt jump in the global metric occurs, but the fundamental nature of the system changes.



(a) (b)

Figure 5: Phase diagram of the Ising model shown in (a) three and (b) two dimensions. The black dot denotes the critical point at which a continuous phase transition occurs with respect to temperature, beyond which a discontinuous phase transition no longer occurs with respect to magnetic field.

Remarkably, the phase diagram of the Ising model shows both visual and conceptual similarities to that of the residential buildings in Milpitas (Figure 4). This correspondence suggests that the collective behaviors identified in Section 3 are indeed phase transitions that occur under varying earthquake magnitudes and uncertainty levels. The analogy between portfolios and systems studied in statistical physics offers significant insights into the understanding of regional seismic risk, including:

- At the critical point, a system exhibits scale invariance [3]; viewing the system at different scales yields statistically identical results. In the context of regional seismic risk analysis, this scale invariance may suggest the potential effectiveness of uniform risk management strategies across decision-making entities of different sizes.
- The critical temperature where the system reaches the critical point can be approximately calculated using a statistical physics method known as the *Landau free energy approximation* [3]. Applying this method to regional seismic risk analysis enables quantification of the critical uncertainty level, providing a clearer understanding of how to assess regional seismic risk within a reliable range and avoid fundamental misrepresentation.
- At the discontinuous phase transition, the *susceptibility*, defined as the sensitivity of the system's response to changes in external excitation, diverges [3]. This phenomenon may imply a sharp decline in a global resilience metric resulting from a slight increase in earthquake intensity, manifesting the abrupt safe-to-failure collective transition in Figures 2a and 2b.

5 CONCLUSIONS

This study demonstrates that regional seismic responses exhibit coherent macroscopic patterns and distinct phase transitions, which are not apparent in bottom-up, component- and scenario-specific simulations. By augmenting a probabilistic regional-scale framework with an epistemic-uncertainty term that reflects model incompleteness, we showed that two control variables, i.e., earthquake magnitude and uncertainty, drive discontinuous and continuous transitions in system-level risk metrics. The discontinuous transition manifests as an abrupt shift from a collective safe state to a collective failure state, whereas the continuous transition emerges as progressive, yet fundamental, changes in the underlying statistics of global performance measures as uncertainty grows. The resulting phase diagram summarizes these phenomena in a compact, intuitive map in hazard–uncertainty space.

Interpreting these phenomena through the lens of the Ising model clarifies the roles of correlation, heterogeneity, and epistemic uncertainty in shaping systemic fragility. Correlation between building capacities and demands is a double-edged sword: it buffers losses under weak hazard intensity yet triggers brittle, system-wide failure once a critical magnitude is surpassed. Increasing the diversity of the building portfolio, and thereby reducing the correlation of damage states, may be more effective in preventing brittle, system-wide failure than incrementally upgrading individual buildings. Moreover, increasing epistemic uncertainty blurs the phase boundary between collective safe and collective failure states, underscoring the importance of a comprehensive modeling framework to prevent misrepresentation of regional risk.

Beyond the specific case study, our findings highlight the value of incorporating statistical-physics concepts, such as phase diagrams and critical phenomena, into performance-based earthquake engineering. These tools provide rapid, top-down insights into how local details scale up to regional consequences, complementing detailed simulations. Future work will (1) incorporate physics-based ground-motion and structural simulations to examine the generality of the observed collective behaviors and (2) explore renormalization group techniques, methods from statistical physics conceived to uncover universal patterns from microscopic details across scales, to identify scale-invariant mitigation strategies that are effective across diverse decision-making entities. Integrating a top-down statistical-physics perspective with conventional agent-based simulations offers a complementary and comprehensive pathway toward enhancing the seismic resilience of communities.

REFERENCES

[1] Heresi, P., and Miranda, E. RPBEE: Performance-based earthquake engineering on a regional scale. *Earthq. Spectra* (2023) **39**(3):1328–1351.

[2] Du, A., Wang, X., Xie, Y., and Dong, Y. Regional seismic risk and resilience assessment: Methodological development, applicability, and future research needs–an earthquake engineering perspective. *Reliab. Eng. Syst. Saf.* (2023) **233**:109104.

[3] Cardy, J. *Scaling and renormalization in statistical physics*. Cambridge University Press, 1996.

[4] Oh, S., Yi, S., and Wang, Z. Long-range Ising model for regional-scale seismic risk analysis. *Earthq. Eng. Struct. Dyn.* (2024) **53**(12):3904–3923.

[5] Jayaram, N., and Baker, J.W. Correlation model for spatially distributed ground-motion intensities. *Earthq. Eng. Struct. Dyn.* (2009) **38**(15):1687–1708.

[6] Rincon, R., and Padgett, J.E. Bias-based comparison of sub-model fidelity and its compounded effect on multiscale infrastructure resilience estimates. *J. Eng. Mech.* (2025) (In Review).

[7] Douglas, J. *Ground-motion prediction equations 1964–2010*. PEER Report 2011/102, Pacific Earthquake Engineering Research Center, University of California, Berkeley, 2011.

[8] Vargas-Alzate, Y.F., Hurtado, J.E., and Pujades, L.G. New insights into the relationship between seismic intensity measures and nonlinear structural response. *Bull. Earthq. Eng.* (2022) **20**(5):2329–2365.

[9] Kang, C., Kwon, O.S., and Song, J. Quantifying uncertainties and correlations of engineering demand parameters of building structures for regional seismic loss assessment. *Earthq. Eng. Struct. Dyn.* (2022) **51**(7):1751–1769.

[10] Zhang, W., Crempien, J.G., Kurtulus, A., Chen, P.Y., Arduino, P., and Taciroglu, E. A suite of broadband physics-based ground motion simulations for the Istanbul region. *Earthq. Eng. Struct. Dyn.* (2023) **52**(4):1161–1181.

[11] Atkinson, G.M., and Assatourians, K. Implementation and validation of EXSIM (a stochastic finite-fault ground-motion simulation algorithm) on the SCEC broadband platform. *Seismol. Res. Lett.* (2014) **86**(1):48–60.

[12] Atkinson, G.M., and Boore, D.M. Ground-motion relations for eastern North America. *Bull. Seismol. Soc. Am.* (1995) **85**(1):17–30.

[13] Federal Emergency Management Agency (FEMA). *HAZUS-MH 5.1 technical manual*. Federal Emergency Management Agency, 2020.

[14] Federal Emergency Management Agency (FEMA). *Seismic performance assessment of buildings, Volume 1: Methodology*, FEMA P-58-1. Federal Emergency Management Agency, 2018.

[15] Rincon, R., and Padgett, J.E. Fragility modeling practices and their implications on risk and resilience analysis: From the structure to the network scale. *Earthq. Spectra* (2024) **40**(1):647–673.

[16] Baker, J.W. Introducing correlation among fragility functions for multiple components. In *Proc. 14th World Conf. Earthq. Eng.* (2008) Vol. 8.

[17] Xiang, M., Shen, J., Xu, Z., and Chen, J. Structure-to-structure seismic damage correlation model. *Earthq. Eng. Struct. Dyn.* (2024) **53**(10):3205–3229.

[18] Baker, J.W., Almeter, E., Cook, D., Liel, A.B., and Haselton, C. A model for partially dependent component damage fragilities in seismic risk analysis. *Earthq. Spectra* (2024) **40**(1):609–628.

[19] Cover, T.M., and Thomas, J.A. *Elements of information theory*, 2nd ed. John Wiley & Sons, 2006.

[20] Zsarnóczay, A., Elhaddad, W., Cetiner, B., Zhong, K., McKenna, F., and Deierlein, G. *SimCenter earthquake testbed: San Francisco, CA*. DesignSafe-CI (2023). https://doi.org/10.17603/ds2-z5sx-zh27.