A SIMPLE COARSE-GRAINED MODEL TO PREDICT INFECTIOUS DISEASE OUTBREAK STATISTICS AND THE CONSEQUENCES OF GLYCOLYSIS INHIBITION ON CELL JUNCTION MECHANICS

BY

GREGORY JOSEPH SCHWARZ

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biophysics and Quantitative Biology
in the Graduate College of the
University of Illinois Urbana-Champaign. 2024

Urbana, Illinois

Doctoral Committee:

Professor Yann Chemla, Chair
Professor Karin Dahmen, Director of Research
Professor Martin Gruebele
Assistant Professor Sangjin Kim

ABSTRACT

Preventative medicine is the best way to keep individuals healthy and the cost of healthcare down. Thus, time and resources can be allocated elsewhere in the medical field if it is possible to prevent ailments that can be used for unpreventable outcomes. Sadly, viral infections are inevitable but occur randomly over a single lifetime. However, when averaging over a larger population, the long-term projected effects of viral infections can be conjectured. Thus, a deeper understanding of stochastic processes and systems can be helpful for diagnostic medicine and virology. This dissertation first highlights the effects of vascular viral infections have on tissue permeability in lung capillaries due to cell's mechanical ineptitudes when deprived of ATP in Chapter I. Then, it focuses on a statistical mean field model of host population dynamics in Chapter II. Chapter III focuses on the statistics of viral and host population dynamics from Lotka-Volterra differential equations, and the application of the model with global disease statistics. Finally, Chapter IV highlights how the flu and COVID-19 fit the model in Chapter II.

ACKNOWLEDGMENTS

I would like to acknowledge my friends and family who helped me throughout my graduate career. I have met wonderful people in various ways. The family I had during my mechano-biology research. I cherished the times we came together through ping pong in the breakroom, late night experiments, dinners, and Oktoberfest. You all got me up to speed on cell biology that I needed coming from a theoretical physics background. I'd also like to acknowledge the friends I have in the Crackling Group that helped me get up to speed and graduate quickly along with our advisor Professor Karin Dahmen. Professor Dahmen gave me the confidence to believe in myself and made me feel like I could accomplish anything. Thank you for your dedication to me and for not allowing me to give up on my degree.

I partook in lots of extracurriculars to keep myself sane outside of my research. Thank you to the wonderful ultimate frisbee community I had through my time playing with Boomland, Illinois Rise, and Prion. Thank you to the encouraging bouldering community at Urbana Boulders that kept me striving to reach my goals. Lastly, a special thank you to the friends I have and the adventure we have been on. (Mike, Jess, Michele, Nicole, Andrew, Felina, Titus, Caitlyn, Nancy, Austin, Rebecca, Dayna, Sayo, Ahsley, and my friends at I-life).

A special acknowledgment to my friend David Bostwick who is no longer with us after a hard-fought battle with cancer. He was a great friend who strived to be better in everything he did. He made sure his friends were loved and that he was there for his brothers and sisters in I-life. He was a bright light during my graduate career, and I deeply miss him.

Dedicated to Jonathan & Ashley Schwarz along with their son

TABLE OF CONTENTS

CHAPTER 1: Glycolysis Inhibition Regulates Endothelial Junctions by Perturbing Actin and Focal Adhesions.....1

CHAPTER 2: A mean field model of host population dynamics model.....21

CHAPTER 3: Fitting Lotka-Volterra host population dynamics with the mean field…….…….…41

CHAPTER 4: The eerie agreement of flu and COVID-19 outbreaks with the model.....53

APPENDIX A: Recovery of HLMVEC junction strength after glycolysis is restored…….…….…63

APPENDIX B: Materials and Methods from Chapter 1…….……...…….66

APPENDIX C: Complementary cumulative distribution plot of all spacer acquisition probabilities for chapter 3…….……..…….……73

References…….....74

Chapter 1: Glycolysis Inhibition Regulates Endothelial Junctions by Perturbing Actin and Focal Adhesions[1][2]

Introduction

Vascular barrier function is a tightly controlled, yet dynamic process that is controlled by adherens junctions, which link neighboring endothelial cells. These junctions are comprised of the surface adhesion molecule vascular endothelial cadherin (VE-Cadherin or Cadherin-5), as well as by catenins, which link the junction complexes intracellularly to the actin cytoskeleton (Cao et al., J Cell Sci, 2019, Le Duc et al., 2010, Tabdili et al., 2011, Buckley et al., 2014). Through this association with actin, the junction complex is maintained under a constant state of tension, which regulates barrier function (Daneshjou N et al, J Cell Bio 2015, Juettner V et al, J Cell Bio 2019, Barry et al., 2014, Huynh et al., 2011). The cells must also respond to external sources of force, such as sheer stress and stretching, which drive remodeling of the junction complexes and associated actin (Salvi et al, Curr Opin Cell Biol, 2018, Tzima et al., 2005, Fang et al., 2019, Gawlak et al., 2014, Ito et al., 2017, Miyake et al ., 2006). These active processes are heavily ATP

[1] Glycolysis Inhibition Regulates Endothelial Junctions by Perturbing Actin and Focal Adhesions https://www.biorxiv.org/content/10.1101/2020.10.22.351148v1.full.pdf

[2] Parallel Activation of Src and Hif1α Increases Localized Glycolytic ATP Generation for Re-assembly of Endothelial Adherens Junctions https://www.biorxiv.org/content/10.1101/2022.11.11.516184v1

dependent, yet little is known about how endothelial cells derive the energy to maintain their barrier function.

In epithelial cells, this energy need is met by increased glucose oxidation. External force is sensed by E-Cadherin based adherens junctions, which then activate AMPK, leading to increased ATP production through mitochondrial oxidative phosphorylation. This process provides the energy needed for cytoskeletal rearrangements and adherens junction stabilization (Bays et al, Nat Cell Bio, 2017). However, unlike epithelial cells, endothelial cells rely primarily on glycolysis, and not oxidative phosphorylation, for their energy needs (De Bock et al, Cell 2013). In particular, the glycolysis regulatory enzyme 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3 (PFKFB3), is a key regulator of endothelial metabolism (De Bock et al, Cell 2013). We thus hypothesized that PFKFB3-driven glycolysis was the source of the ATP required to maintain the integrity of endothelial adherens junctions. The maintenance of tight endothelial junctions is a critical function of the vascular endothelium, because it prevents fluid and protein leak from the blood into the tissue. During pathological processes such as inflammatory lung injury or diabetic retinopathy, this vascular integrity is compromised, and the resulting tissue edema is a key determinant of disease progression (Komarova Y, Circ Res 2017).

In this chapter, we investigated the impact of PFKFB3 inhibition on VE-cadherin-mediated contacts between primary human endothelial cells. Traction force imbalance (TFIB) measurements and the use of a FRET-based VE-cadherin tension sensor (Conway et al, Curr Biol, 2014) revealed mechanical changes at the level of both cell-cell junctions and adherens junction proteins. Using PFK15, which is a pharmacological inhibitor of PFKFB3 (Clem et al, Mol Cancer Ther, 2014), we show that PFK15 treatment significantly decreased both the overall junctional tension and the local force on VE-cadherin molecules. Importantly, these mechanical changes were not confined to local

perturbations at cell-cell contacts but reflected global changes in force distributions throughout the cell, including the disruption of focal adhesions and loss of cytoskeletal coherence. These findings reveal the importance of glycolysis in maintaining global cell mechanics, which in turn regulates barrier homeostasis in endothelial cells.

Results

Cell adhesion on microcontact printed fibronectin patterns on 40kPa poly (acrylamide) gels:

Measurements of the force per unit length (tension) at junctions between human lung microvascular endothelial cells (HLMVECs) were carried out with cells on arrays of microcontact-printed fibronectin on polyacrylamide (PAA) gels. The stamped fibronectin pattern of two, tangential 1600 μm^2 circles is shown in Figure 1.1a, and a representative fluorescence image of micropatterned arrays of fluorescently labeled fibronectin is shown in Figure 1.1b.

Cells adhered to the patterns (Fig. 1b.1), but the shapes of cells on these patterns were variable, such that the junction lengths and junction orientations relative to the long axis of the doublet differed. We determined the junction tension from the average traction stress exerted by each cell, but cell shapes can affect the traction generation. We therefore conducted sufficient traction force imbalance (TFIMB) measurements to ensure that we accounted for the variability in cell shapes and could thus determine the tension differences between control cells and those with inhibition of glycolysis via the PFKFB3 inhibitor PFK15.

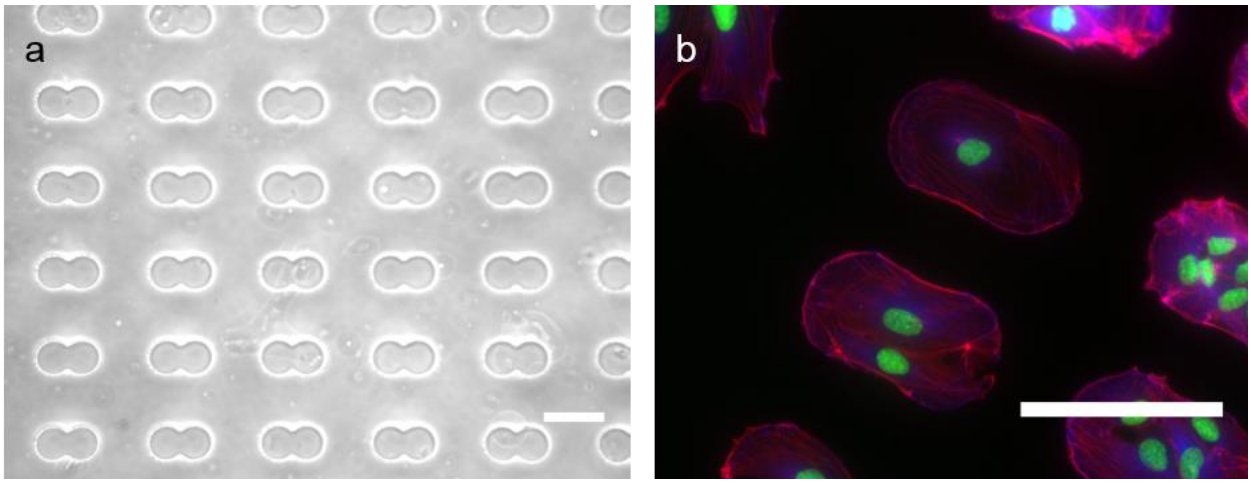


Figure 1.1 Cell doublets on micro contact printed fibronectin patterns on polyacrylamide gels: a) PDMS stamp pattern of tangential 1,600 μ^2 circles. Scale bar = 90 μm. b) Cell doublets on fibronectin patterns printed on 40kPa gels. Cells were immunostained-for paxillin (blue), actin (red), and DAPI (green).- Cells were imaged at 40x magnification with a scanning confocal microscope. Scale bar = 45 μm.

Inhibiting glycolysis reduces force on VE-cadherin adhesions:

Treatment with PFK15 correlated with a decrease in the force on VE-cadherin bonds. Studies compared the CFP/FRET ratio of the VE-cadTS at junctions between patterned HLMVECs, in the absence and presence of PFK15. Figure 1.2 shows the fluorescence images of untreated cells (Fig. 1.2a) and of cells 40 min after PFK15 treatment (Fig. 1.2b). As determined from the normalized CFP/FRET ratios (Fig. 1.2c), glycolysis inhibition reduced the normalized CFP/FRET ratio from 0.97 ± 0.06 to 0.65 ± 0.03, indicative of a loss in force on VE-cadherin complexes. This difference was statistically significant ($p < 0.001$, $N_C = 29$, $N_I = 34$).

We also observed morphological differences. Junctions between untreated cells were broad and typical of the overlapping contacts between endothelial cells. Although the width of the junction between the PFK15 treated cells in Fig. 1.2b appear narrower and straighter, this apparent difference was not observed uniformly across cells. Fig. 1.2b also shows some membrane

invaginations at junctions between PFK15 treated cells that were observed more frequently than with treated cells.

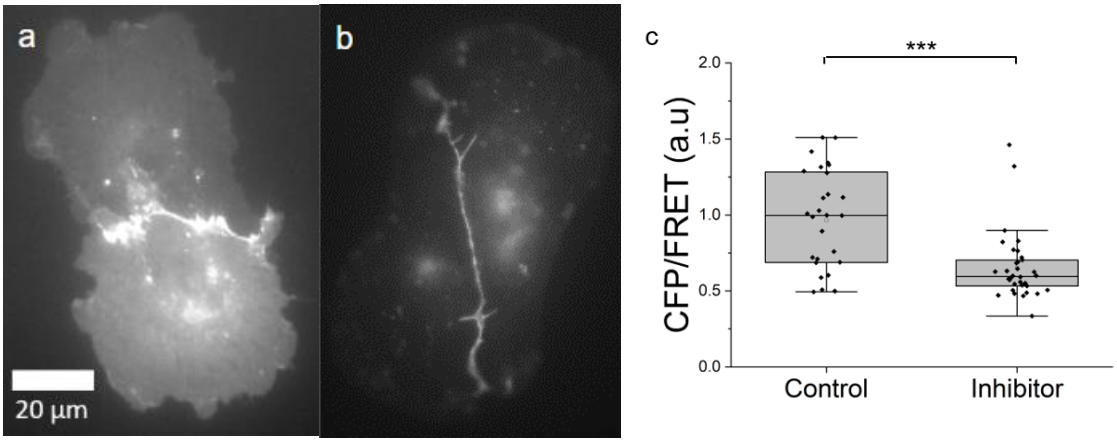


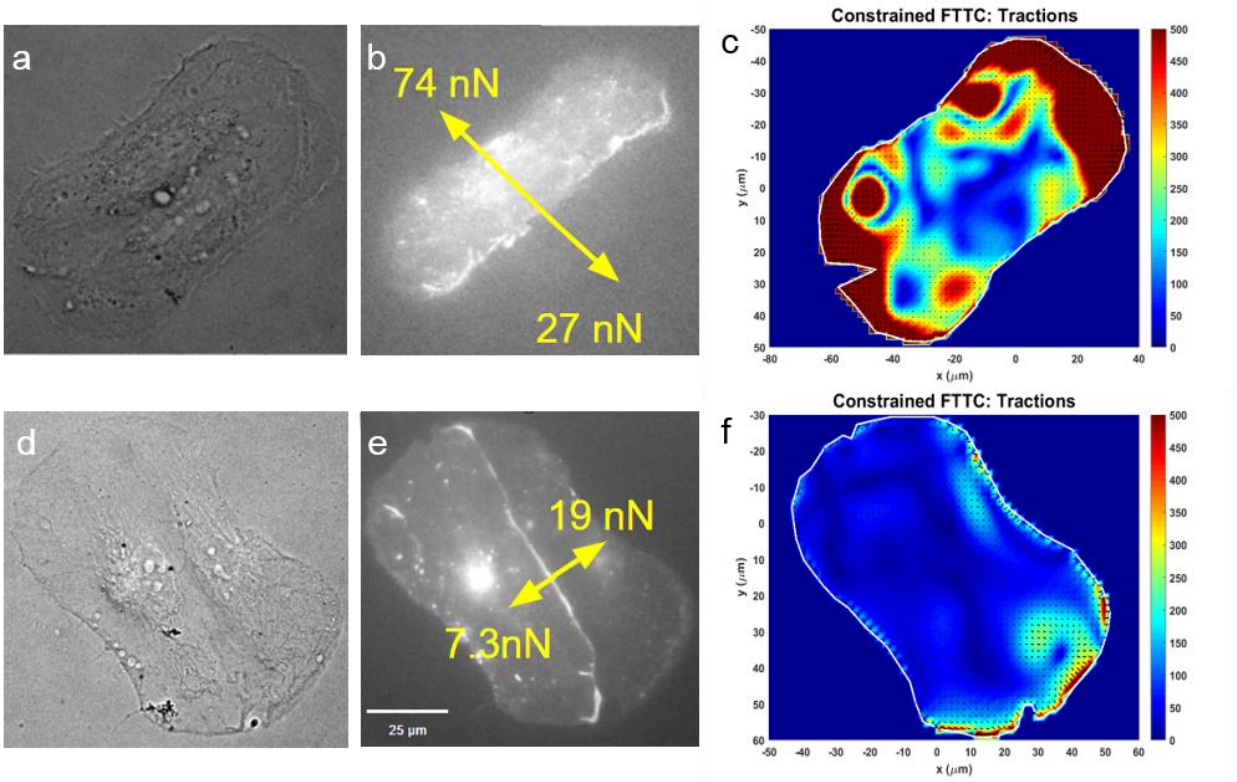
Figure 1.2: PFK15 treatment reduces the tension on VE-cadherin mediated junctions between HLMVEC pairs. a) Fluorescence images of the VE-cadTS at the junctions between a) control and b) PFK15 treated HLMVECs expressing VE-cadTS (images obtained with the YFP channel were enhanced for presentation purposes only). c) CFP/FRET ratio measured in the region of interest at cell-cell junctions. Average CFP/FRET is significantly lower in the PFK15 treated cells than in the control (***$p < 0.001$, $N_C = 29$, $N_I = 34$).

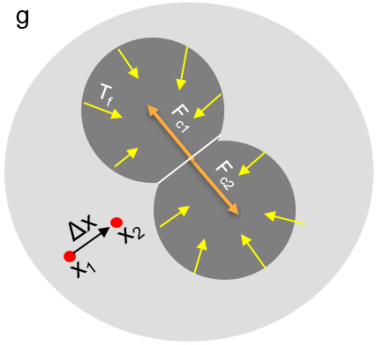
PFK15 treatment alters cell tractions and tension at cell-cell junctions:

To determine the effects of glycolysis inhibition on the average junction tension, which depends on global cell mechanics as well as on local changes at intercellular adhesions, we performed Traction Force IMBalance measurements (TFIMB). Measurements compared the tension at the junctions, with and without PFK15 treatment. Figures 1.3a&d are DIC images of HLMVEC doublets on a 40 kPa gel, at 40 min without (Fig. 1.3a) and with (Fig. 1.3d) PFK15 treatment. Figures 1,3b&e compare the traction stresses exerted by untreated and inhibitor-treated cells, respectively. Comparison of Figs. 1.3b&e shows that PFK15 treatment decreases the traction stresses at the basal plane, relative to untreated controls (Fig. 1.3b). This is apparent in the red (high stress) to blue (low stress) shifts in the heat maps.

Figures 1.3c&f display the net force on the junction produced by each cell in the doublet. Junctions were visualized with GFP-VE-cadherin. Figure 1.3g illustrates the force diagram underlying the TFIMB measurements. Here, displacements of the beads in the gels are split into their x and y components, which define the x and y components of the traction stresses. The latter are used to determine the net force orthogonal to the junction. The difference in traction stresses is balanced by the net force on the junction (Maruthamuthu et al., 2011). The average junction tension (Fig. 1.3g) is the net force divided by the junction length, which is visualized with GFP-VE-cadherin.

In the absence of PFK15 (Fig. 1.3c), force balance analysis indicated an average net positive force on the junctions. As shown in Fig. 1.3f, PFK15 treatment decreases the magnitudes of the net force exerted by each cell, with a corresponding decrease in the junction tension, compared to untreated cells (Fig. 1.3c). Quantitatively, the average tension decreased from 1.7 ± 0.4 to 0.3 ± 0.1 nN/μm. This statistically significant difference ($p < 0.01$, $N_C = 20$, $N_I = 14$) is qualitatively consistent with the VE-cadherin tension sensor results (Fig. 1.2). This result reveals that PFK15-dependent changes in global cell mechanics reduces the inter-endothelial tension.





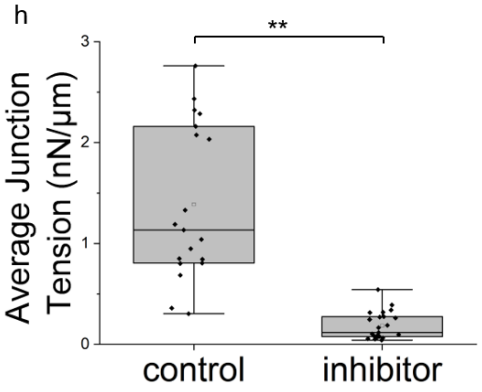


Figure 1.3: Traction Force Imbalance (TFIMB) measurements of HLMVEC pairs. Substrates were 40 kPa gels. DIC Images of untreated (a) and PFK15 treated (c) cell pairs on fibronectin patterns. Fluorescence image of control (b) and PFK15 treated (e) cells expressing GFP-VE-Cadherin and the determined forces exerted on the junction by each cell. Heat maps of traction stresses exerted by (c) control and (f) PFK15 treated HLMVEC doublets seeded on stamped fibronectin patterns of tangential circles. (g)

Figure 1.3 (cont.): Force diagram of HLMVEC doublets on a substrate showing the traction force vectors relative to the axis of the junction. Tr is the traction force, F_{c1} and F_{c2} are the forces exerted by cell 1 and cell 2, respectively, and Dx is the length of the junction. (h) Average junction tension of control and PFK15 treated cells ($p < 0.01$, $N_C = 20$, $N_I = 24$).

PFK15 treatment alters the magnitudes and orientations of traction stress vectors:

PFK15 disrupted the magnitude and orientational coherence of the individual traction vectors exerted by the cells. The direction and magnitude of the forces determine the junction tension, so we investigated the distribution of vector orientations, with and without PFK15 treatment. Figures 1.4a&b show wind rose plots of the angles of the traction stresses relative to the junction normal, in the absence and presence of inhibitor, respectively. In untreated cells (Fig. 1.4a), the absolute values of the angles between the traction force vectors and the junction normal are primarily between 0^o and 35^o such that the traction vectors are largely oriented perpendicular to the junctions. By contrast, the stress vectors exerted by PFK15 treated cells (Fig 1.4b) appear to be more randomly distributed.

Inhibiting glycolysis also altered the distribution of the magnitudes of the traction stresses. This is shown in the histograms in Figure 1.4c. In the untreated cells, the most probable traction stress is ~200Pa, and there is a broad, long tail. With PFK15 treated cells, the most probable traction stress decreased to ~50Pa, the distribution is narrower, and the tail is substantially shorter. Taken together, these results show that inhibiting glycolysis disrupts both the relatively uniform orientation traction stresses normal to the junctions (Figs. 1.4a,b) and shifts the distribution of traction stresses to smaller values (Fig. 1.4c).

We confirmed the impact of PFK15 treatment on global cell mechanics, independent of cell-cell junctions, by analyzing traction stresses of individual cells, using Traction Force Microscopy (TFM). Figure 1.5 compares traction stress heat maps of individual HLMVECs on 40

kPa gels without (Fig. 1.5a) and after (Fig. 1.5b) PFK15 treatment. In untreated cells, the force vectors are oriented radially inward (Fig. 1.5a) and actin fibers span the width of the cell, as reported previously (Cai et al., 2010). PFK15 treatment resulted in more randomly oriented stress vectors (Fig. 1.5b). There was also a drop in the magnitude of the root mean square (RMS) traction stress exerted by individual cells from 210 Pa ± 20 to 100 ± 10 Pa ($p < 0.001$, $N_C = 28$, $N_I = 29$).

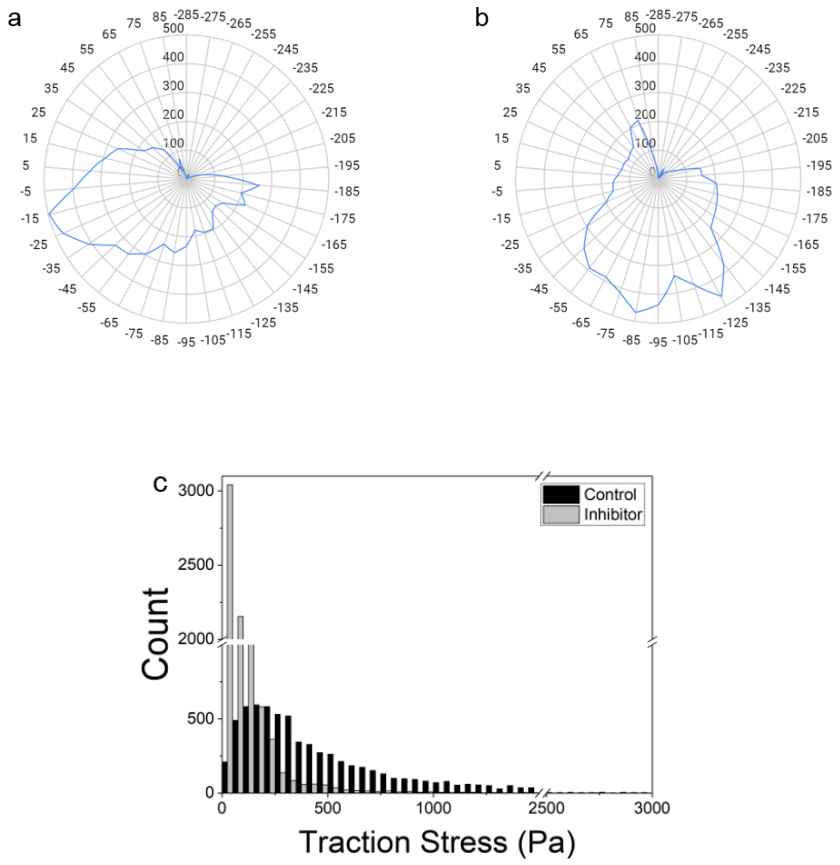


Figure 1.4: Traction stress distribution and orientation. Wind rose plots of the angles of traction force vectors relative to the junction normal in (a) control and (b) treated HLMVEC doublets. (c) Histogram of the magnitudes of measured traction stresses exerted by cells doublets on printed fibronectin patterns on 40 kPa gels. Black bars indicate control cells ($N_C = 20$) and gray bars were obtained with inhibitor treated cells ($N_I = 24$).

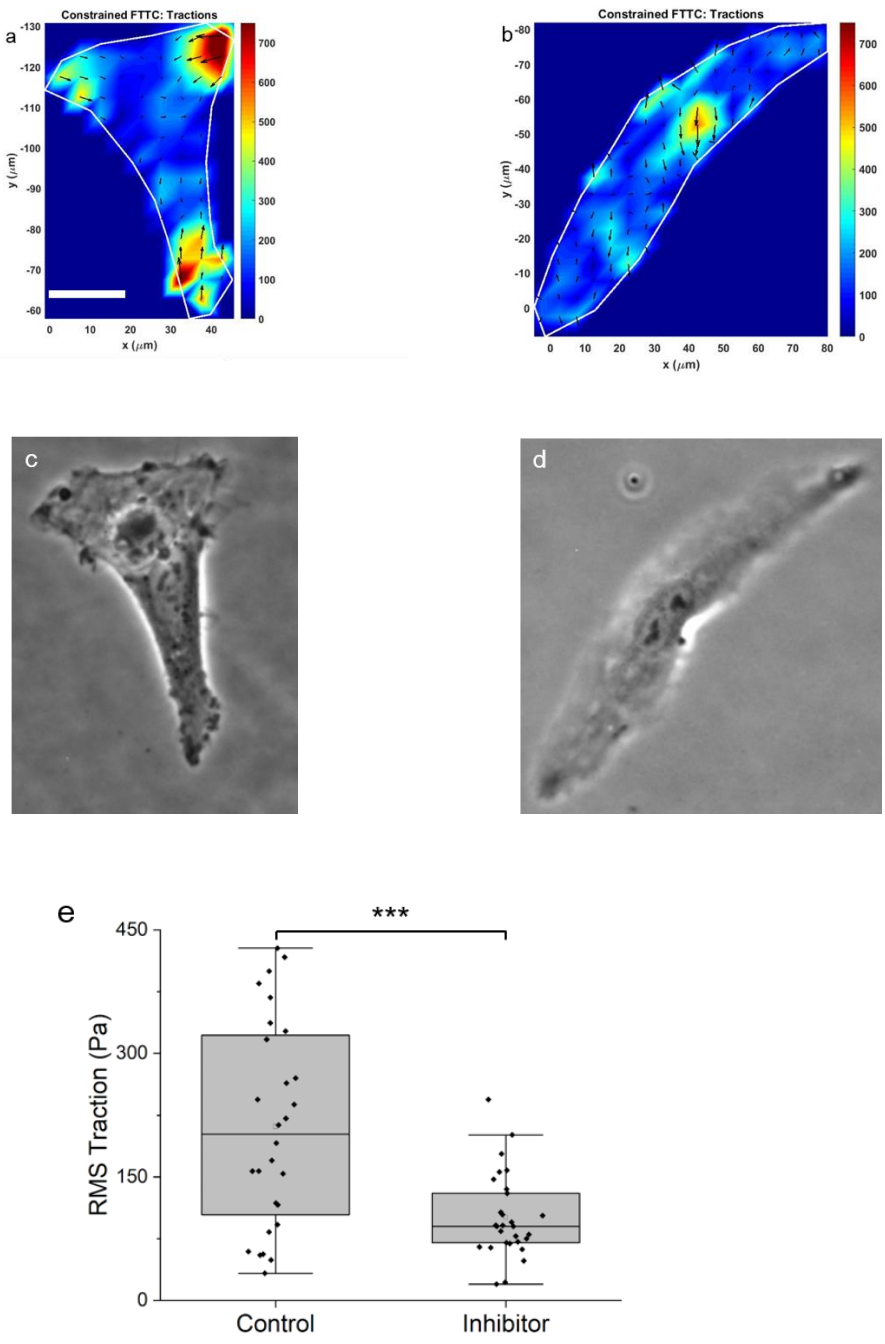


Figure 1.5: PFK15 treatment reduces the traction force amplitude and orientational uniformity. a) Traction force heat maps of treated and b) PFK15 treated, single HLMVECs. Cells were seeded at low density on fibronectin coated 40kPa gels. DIC images of c) control and d) PFK15 treated HLMVECs. e) Root mean square traction stress (Pa) generated by HLMVECs 40 min after

Figure 1.5 (cont.): PFK15 treatment, compared with untreated controls. Cells were seeded on the substrates described in parts a and b ($p < 0.001$, $N_C = 28$, $N_I = 29$). Scale bar is 20 μm.

Glycolysis inhibition disrupts actin networks and focal adhesions:

To investigate subcellular changes underlying the decreased magnitude and orientational coherence of the traction stress vectors, following PFK15 treatment, we imaged F-actin and focal adhesions. DAPI staining visualized the nuclei of the individual cells, and GFP-VE-cadherin visualized the cell-cell junctions. In untreated cells, paxillin staining of focal adhesions at the basal plane was distinct and punctate (Fig. 1.6a), with larger focal adhesions at the cell perimeter. Paxillin was also depleted in the region adjacent to cell-cell junctions. Actin stress fibers were prominent at the basal plane, and crisp F-actin fibers spanned the length of the cell (Fig. 1.6b). Dense cortical F-actin staining was also apparent parallel to the cell-cell junctions. Vector maps of the orientations of actin fibers revealed the relatively uniform fiber alignment, relative to junctions, in untreated cells (Fig. 1.6c).

PFK15 treatment disrupted both focal adhesions and F-actin. Paxillin staining was more diffuse (Fig. 1.6d), and puncta at the basal plane were fewer and smaller. F-actin was also more diffuse throughout the cell (Fig. 1.6), and there were fewer, well-defined actin fibers. The vector maps (Fig. 1.6d) of actin fibers in treated cells also revealed more random fiber orientations.

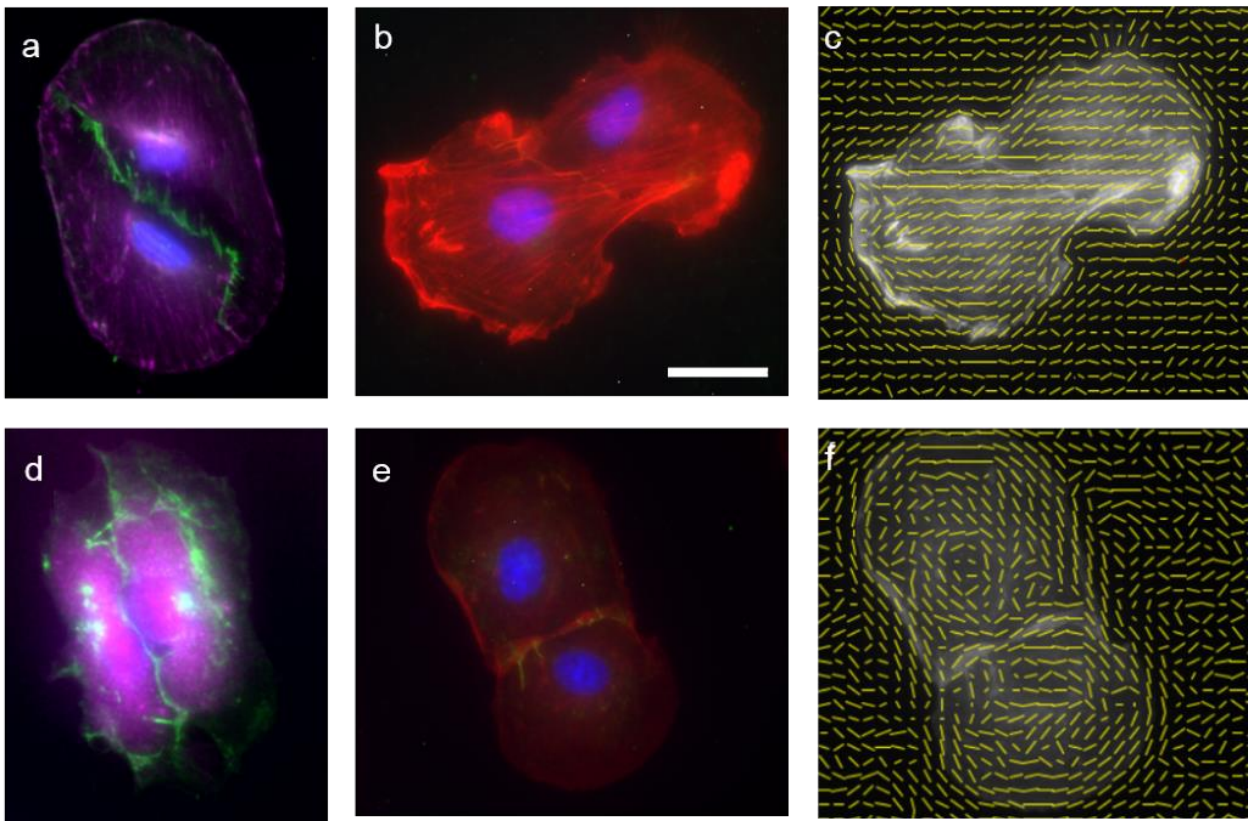


Figure 1.6: Inhibiting glycolysis disrupts focal adhesions and F-actin: Merged immunofluorescence images of paxillin (purple) and VE-cadherin (green) in a) untreated and d) PFK15 treated HLMVEC doubles. Merged immunofluorescence images of actin (purple) and VE-cadherin (green) in b) untreated and e) PFK15 treated HLMVEC doublets. Nuclei are stained with DAPI. c/f) Vector maps showing the directionality of the actin fibers in c) untreated and f) PFK15 treated HLMVECs. The vector maps correspond to cells in panels b & e. In all images, cells were on micro contact printed fibronectin on 40 kPa gels. Scale bar = 25 µm.

Glycolysis inhibition disrupts F-actin fibers and reduces fiber anisotropy:

To visualize the change in the actin fibers between the PFK15 treated and untreated cells, we used ImageJ's "plot profile" tool to generate intensity scans associated with the F-actin across the cell. The scan direction was orthogonal to the majority of visible actin fibers. We examined the phalloidin intensity profiles (Fig. 1.6 a&c) of both PFK15 treated and untreated cells across the nucleus. In scans of untreated cells, there are distinct intensity peaks (Fig 1.6b) and valleys associated with the fibers. In PFK15 treated cells, the fibers aren't as distinct: scans contain fewer

of the larger peaks and valleys (Fig. 1.6d). The large drop in actin intensity across the nuclei also revealed a significant reduction in perinuclear actin.

To characterize the difference in actin fiber alignment in treated versus untreated cells, we used the ImageJ plugin "Fibriltool" (Boudaoud et al., 2014), which calculates the fiber anisotropy and a directionality vector within a region of interest (ROI). Here the ROI is the entire cell. Figure 1.7 shows the cell doublets and the ROI (yellow outline), as well as the magnitude and direction of the calculated actin orientational anisotropy (red line). In untreated cells, the fiber orientation is anisotropic, with the majority of fibers directed along the long axis of the cell doublets and nearly perpendicular to the junctions. The calculated anisotropy value for the untreated cells is significantly larger than that of PFK15 treated cells, as indicated by the relative lengths of the red lines. The anisotropy number for actin in PFK15 treated cells decreased from 0.24 ± .09 (SEM) to 0.013 ± .012 (SEM) (Fig. 1.7; $N_{control} = 46$, $N_{inhibiter} = 33$, $p < 0.005$).

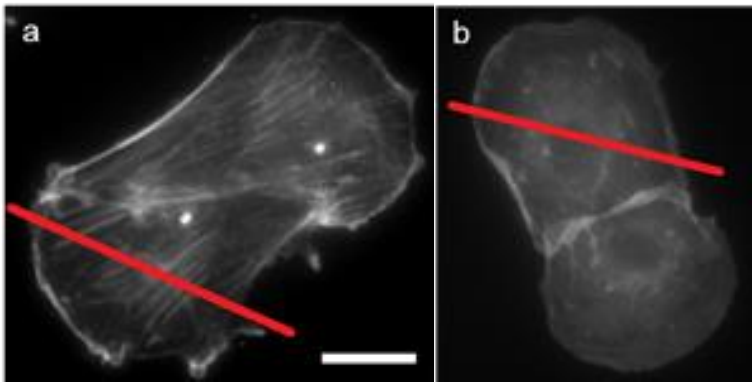


Figure 1.7: PFK15 alters the width and length of F-actin fibers.

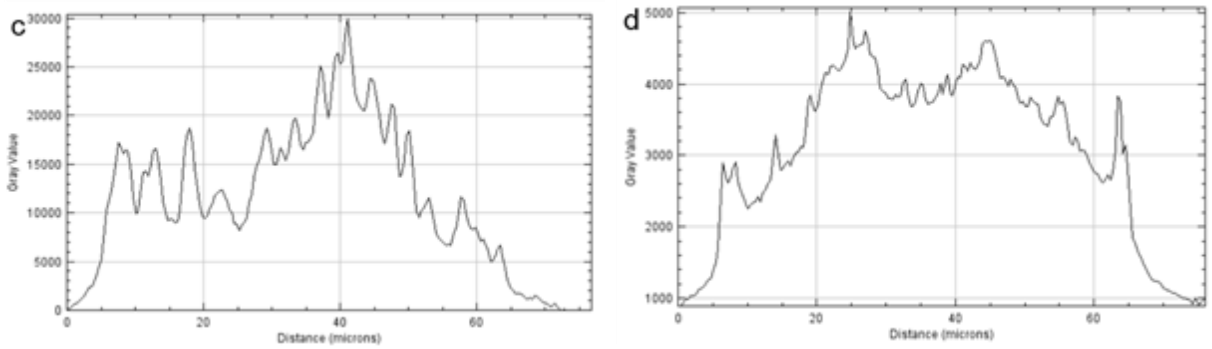


Figure 1.7 (cont.); Immunofluorescence images of F-actin in HLMVECs in a) controls and b) after PFK15 treatment. The red line indicates the showing where intensity scans were taken. c) Intensity scans across the cell in panel a and d) intensity scan across the cell in panel b. F-actin intensity scans in HLMVECs show distinct peaks corresponding to F-actin fibers whereas the actin is more diffuse and the fibers are narrower and shorter in PFK15 treated cells. Cells were seeded on fibronectin patterns on 40 kpa gels. Scale bar = 25 µm.

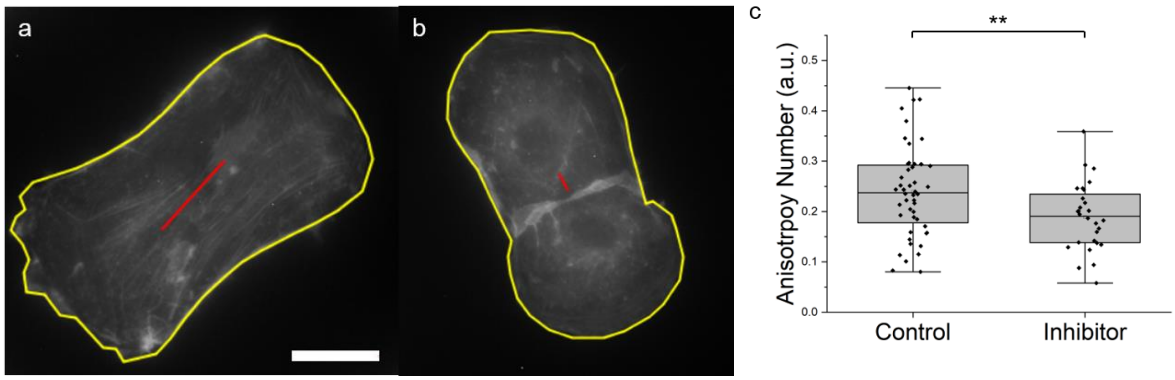


Figure 1.8: Anisotropy of F-actin fibers decreases following PFK15 treatment. The anisotropy orientation and magnitude corresponding to F-actin fibers within the cell boundary (yellow outline). Cells were seeded on microcontact printed fibronectin patterns on 40 kPa gels. Phalloidin-stained actin was imaged at 60x magnification. Quantified actin anisotropy is shown for a) untreated controls and for b) PFK15 treated HLMVECs (generated by Fibriltool). c) Average anisotropy number of actin fibers in untreated and treated cells. The difference is statistically significant (**$p < 0.002$, $N_C = 46$, $N_I = 33$). Scale bar = 25 µm.

Inhibiting glycolysis reduces the diffusivity and mobile fraction of junctional VE-cadherin:

Actin disruption would also affect local processes at inter cellular adhesions, such as VE-cadherin recycling and clustering that regulate intercellular adhesion (Cao et al., 2019). We thus used fluorescence photobleaching after recovery (FRAP) measurements to determine whether observed changes in actin and energy depletion also results in local changes at junctions. Here, we tested whether PFK15 alters VE-cadherin dynamics. Images in Fig. 1.8a&b show the fluorescence reduction and recovery in control and PFK15 treated cells, respectively. At t=0, the images show the fluorescence reduction, after illuminating a 17 μm^2 region of the junction. Figure 1.8c shows the normalized intensity versus time, after the initial bleach. The data were fitted to an intensity recovery model that assumes anomalous diffusion (Fig. 1.8c). Data were fit to the following equation using non linear least squares regression:

$$\frac{I(t)}{I_o} = A(1 - e^{-(\frac{t}{\tau})^B}) \qquad \text{Eq.1.1}$$

Here, A is the mobile fraction and $\boldsymbol{\tau}$ is the recovery time (Lorén et al., 2015). The exponent, B indicates anomalous sub diffusion when B<1 and Brownian diffusion when B =1.

The fits in Fig. 1.8c show that Eq. 1.1 describes data obtained under both conditions, and B < 1 in both cases. With untreated cells, the fitted parameters for the recovery time and mobile fraction were, respectively 10.9 s ± 0.2 and 0.576 ± 0.002. PFK15 treatment altered both the recovery time and the mobile fraction, with the best-fit parameters being, respectively, 4.7 s ± 0.5 and 0.550 ± 0.008. Exponent B decreased from 0.333 ± .005 to 0.172 ± 0.007. The fluorescence recovery time for the control cells was significantly longer than for PFK15 treated cells ($p < 0.001$),

and the mobile fraction was significantly lower in inhibitor-treated cells ($p < 0.01$). These results show that PFK15 treatment also perturbs local VE-cadherin dynamics.

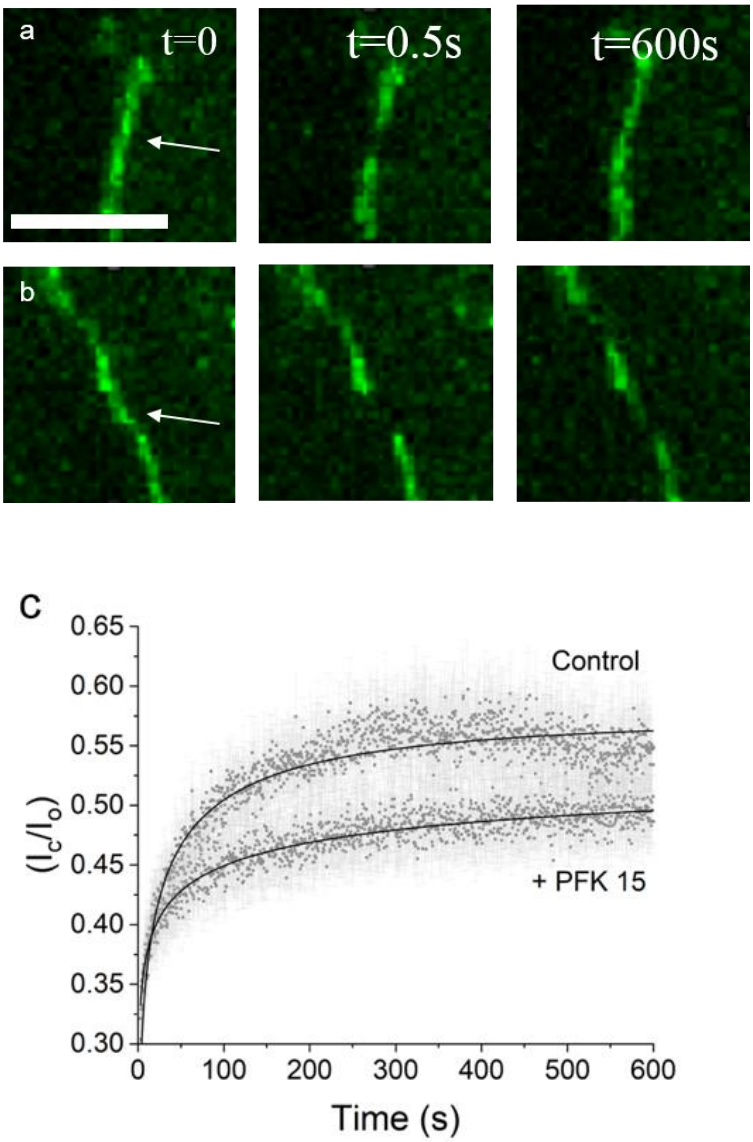


Figure 1.9: PFK15 treatment alters the recovery time and mobile fraction of VE-cadherin at intercellular junctions. Fluorescent images of GFP-VE-cadherin at junctions between HLMVECs in a confluent monolayer on fibronectin-coated, 40kPa gels. The arrows in a) and b) indicate the region of interest prior to bleaching. a) Control and b) PFK15 treated HLMVECs prior to photobleaching. The bleached region is shown for control and PFK15 treated cells initially and after 10 min. c) Weighted, nonlinear least squares fits of the fluorescence recovery data to equation 1. The best fit parameters are given in the text. ($P_{\text{Recovery Time}} < 0.001$, $P_{\text{Mobile Fraction}} < 0.01$, $N_C = 7$, $N_I = 8$). Scale bar = 25 µm.

Discussion

Loss of vascular barrier integrity is observed in multiple diseases such as acute respiratory syndrome (ARDS) or diabetic retinopathy, resulting in the massive influx of proteinaceous fluid into the tissue that compromise organ function (Komarova Y, Circ Res 2017). Much of the research on the mechanisms of endothelial junctional integrity and traction forces has focused on the expression levels of the molecular mediators such as VE-cadherin (Komarova Y, Circ Res 2017), but little is known about the metabolic regulation of junctional forces. The main finding of this study is that impairing glycolysis—the principal source of ATP in endothelial cells (De Bock et al., 2013) —reduces tension at cell-cell junctions, in part, by disrupting actin organization and focal adhesions distal from the cadherin-mediated intercellular junctions. Measurements at the macroscopic level of intercellular contacts and at the level of VE-cadherin proteins revealed that inhibiting the glycolysis regulatory enzyme PFKFB3 resulted in a significant drop in junctional tension. Hypoxic conditions are known to potentiate endothelial barrier disruption (Tojo et al., 2018), but most studies have focused on identifying local biochemical changes that destabilize cell-cell contacts. These results reveal that ATP depletion results in global changes that also impinge on intercellular adhesions (Conradi et al., 2017).

The force balance measurements highlight two parameters that contribute to the reduced force on inter endothelial contacts; the shift in the magnitude of the traction stresses and the loss of orientational coherence of traction vectors relative to the junction normal. The latter is visualized using wind rose plots, which reveal a more random distribution in vector angles relative to the junction normal, following PFK15 treatment. Either change individually would reduce the tension, but PFK15 treatment alters both. Importantly, the change in traction stresses used to determine the junction tension was not simply due to the disruption of intercellular adhesions.

Cells remained in contact throughout the measurements. Additionally, in isolated cells, PFKFB3 inhibition similarly disrupted the radial orientation of traction force vectors and disrupted actin stress fibers that typically span the cell.

The decrease in the force on cell junctions correlates with the lower tensile force on VE-cadherin complexes, which influence endothelial barrier integrity (Komarova et al., 2017). This might be expected. However, in shear sensing, the onset of fluid flow over endothelial monolayers results in a decrease in force on VE-cadherin and an increase in force on platelet endothelial cell adhesion molecule one (PECAM-1) at junctions (Conway et al., 2013). Here we attribute the lower tension on VE-cadherin, in part, to the disruption of actin fibers to which cadherin is mechanically linked. We did not detect an obvious decrease in VE-cadherin at junctions—a change that would likely increase the force on the remaining cadherin bonds. At the same time, lower tension could destabilize VE-cadherin complexes, by decreasing catenin dependent connections to actin (Buckley et al., 2014, Tabdili et al 2011, Le Duc et al., 2010). Prior studies reported causal relationships between force on cadherin complexes and junction stability (Giannotta et al., 2013). These results show that metabolic dysfunction impacts both protein level mechanics and the overall junction tension, by perturbing the actomyosin cytoskeleton and tethering forces in the cell.

Energy depletion induces actin disintegration and gap formation in endothelial monolayers (Kuhne et al., 1993). The immunofluorescence images and F-actin analyses confirm microfilament disruption, but they also highlight broader impacts on global and junctional mechanics. Endothelial junctions are regulated by a balance of tethering and contractile forces that are propagated through the cytoskeleton (Dudek et al., 2001). In single cells, myosin II and actin networks spanning the cell mechanically couple distal focal adhesions (Cai et al., 2010). Here, immunofluorescence images show that cell-cell contacts are organizing centers that orient actin

fibers and traction stresses, which in turn contribute to the force on intercellular adhesions. Actin disintegration severs this connectivity, and thereby decreases the junction tension. At the same time, energy depletion and actin disruption would also reduce contractile forces required to maintain focal adhesions (Bershadsky et al., 2003), accounting for the observed decrease in punctate paxillin staining. Some focal adhesions and actin fibers remain after PFK15 treatment, but the traction generation and mechanical connectivity with cell-cell contacts are significantly impaired.

Tension influences the stability of inter endothelial junctions, but local, ATP-dependent processes such as actin remodeling, as well as cadherin endocytosis, exocytosis and clustering also maintain junctional homeostasis (Wu et al., 2019). PFKFB3 inhibition reduced the fluorescence recovery time (increased mobility), possibly by altering cytoskeletal interactions. Yet, the mobile fraction of VE-cadherin is lower. Although determining the biochemical mechanisms underlying the FRAP results is beyond the scope of this work, the findings confirmed that energy depletion also impinges on local processes at cell junctions. Our findings thus demonstrate that inhibiting glycolysis impacts both local and global processes that regulate adherens junctions.

Conclusions

Energy depletion destabilizes endothelial adhesions, resulting in gap formation and compromised barrier function. The findings reported here show that inhibition of glycolysis not only impacts the tensile force on VE-cadherin complexes but also disrupts the global actin network and distal focal adhesions that regulate global cell mechanics and the tension at cell-cell junctions. Thus, the impact on endothelial junctions due to impaired glycolysis is not limited to compartments near cell-cell contacts, but also involves global changes in actin and focal adhesions distal from junctions. Understanding this link between glycolysis and junction mechanics highlights the

importance of endothelial metabolism in maintaining barriers function during homeostasis. It also suggests that augmenting glycolytic ATP production could be important for restoring junctional integrity and forces during pathological processes.

Chapter 2: A mean field model of host population dynamics model

Introduction

Epidemics are an abnormal number of cases for a disease for a geographical region **(Evans, 2009)**. During the recent COVID epidemic, there were times when many people were infected, separated by times when only a few people were carrying the virus. It is important to understand what determines the sizes of the outbreaks and how they can be reduced or prevented to estimate the needed preparations such as hospital capacities to control, or even stop, an epidemic. A way to model these outbreaks could give the needed understanding. Here we propose a simple model for outbreaks to show that their statistics can be used to infer the underlying disease dynamics and potentially predict the duration of the epidemic. The results also predict how the dynamics should change depending on whether the society is well mixed, with every host potentially in contact with every other host or a host can only be in touch with nearest neighbors. The cohabitation and interactions between the host and virus play a significant role in disease progression due to viral infections.

In an ecosystem, bacteria can be infected by a virus or be immune to a virus via CRISPR- and the Cas-9 proteins **(Labrie et al., 2010, van Houte et al., 2016, Mayer et al., 2016)**. The delicate balance of host immunity against a virus and a virus's ability to mutate to circumvent said immunity **(Sanjuán et., al)** in an ecosystem causes rapid bursts of viral population and drops of host population. If a strain of bacteria is immune to a virus, the strain population will remain near carrying capacity and stifle the viral population growth. If the strain is not immune to a virus, then the strain population will shrink and the virus population will increase (**Pilsof et., al).** The latter can cause virus populations to burst rapidly. These rapid bursts of a virus' population, associated with rapid drops in the bacterial host population, are avalanches (*i.e*. domino-effect like cascades in the dynamics). Previous work **(Sethna et. al., 2001, Mehta et. al., 2002, Salje et., al, 2014)**

shows that a variant of the model that we use to describe epidemics reveals outbreaks of a broad range of sizes near a critical amount of disorder in the immunity in the population. For more homogenous populations the outbreaks can affect a large part of the entire population. For more inhomogeneous immunities in the population, the outbreaks remain small.

The scaling properties of the outbreak statistics are different in the well mixed case (mean field theory) compared to the case that has only nearest neighbor interactions. In this paper, we focus especially on the well-mixed case and discuss how virus population dynamics (or the host population) can be modeled by mean field models with long-range restoring fields. We show that the observed intermittent dynamics are reminiscent of avalanches and crackling noise that has been seen in many other dynamical systems (**Sethna et. al., 2001).**

Theory

A simple mean field model for well mixed cases, that gives statistics for fluctuations in other systems (magnets, earthquakes, neurons), can describe the host population dynamics. Here, we will discuss an innovative way to use the simple mean field model to describe host populations dynamics through two versions of simulations. Both simulations use the mean field model of **equation 1**. The index *i* labels the hosts and can take any value between 1 and the maximum total number (carrying capacity) N of hosts: $i = 1,\ldots,N$. Each host has a local spin (S_i), resulting in either a healthy host ($S_i = -1$, pointing down) or an infected and lysed host ($S_i = 1$, pointing up).

Each host is exposed to an effective dangerous virus environment that varies from site to site that has two contributions: (1) virus mutation into dangerous strains, modeled by a term H(t) below, and (2) increase of virus numbers by lysing of hosts that have been infected, modeled by the positive coupling term $J > 0$ (related to how contagious the disease is). Mathematically J is

related to the number of viruses that are released upon lysing of a host (*i.e.* S_i flips from -1 to +1) multiplied with the probability that another host gets infected by the virus. These two contributions will cause the viral population to grow and the host population to drop. As more spins flip from -1 to 1 (*i.e.* as more hosts lyse) $H_{effective,i}$ grows. $H_{effective,i}$ plays the role of a local effective "field" for host *i*.

To counter the virus attacks a host has two defenses: (1) an innate immune system represented by a random field h_i and (2) the ability to acquire immunity after infection modeled through the negative coupling $J_{inf} < 0$. These defenses result in a shrinking viral population and growth for the host population. Below we give more details on the mathematical description. We represent the local environment by **equation 2.1.** $H_{effective,i}$ initially starts at a sufficiently negative value to model a situation where all hosts are healthy. Through virus mutation into more dangerous strains (with slow mutation rate dH/dt Ω) H(t) is subsequently slowly increased. The rate at which it is increased corresponds to the mutation rate of the viruses.

$$H_{effective,i}(t) \equiv H(t) + \sum_{k=0}^{N}\left(\frac{J}{N}\right)S_k + (h_{i,t} - J_{inf}Q(t)) \qquad \textbf{(2.1)}$$

In the context of models for magnets, the last term in equation 2.1, $I_i(t) \equiv h_{i,t} - J_{inf}Q(t)$ **(Kuntzt et. al., 2000)**, is called the "demagnetizing field" ($H_{demag,i}(t)$). But in our case, it effectively models the induced immunity in the hosts. Here, J_{inf} is the demagnetization constant that is related to the probability that a surviving host boosts its immunity when it is exposed to the virus. Essentially, each time a host gets infected, there is a probability that the host can acquire immunity against the virus via the CRIPSR-Cas9 by acquiring a DNA spacer from the virus and incorporating it into the host's genome. The process using CRISPR-Cas9 is called CRISPR acquired immunity (**Childs et., al. 2012, 2014**). Such bacteria will reproduce and pass on the acquired immunity, causing the host population to develop immunity against the virus. We

have S_i = -1 for each host *i* that has not lysed and becomes less probable to lyse due to the additionally acquired immunity. The hosts with S_i = -1 will have a negative local effective field ($H_{effective,i}$) while the hosts with S_i = 1 have a positive local effective field.

A coarse-grained simulation was created to show how the virus population bursts can be modeled through the above mean field model (**equation 2.1**) The code initializes setting J_{inf} and disorder (R). The disorder is the width of the Gaussian distribution is directly correlated with the Gaussian width. There is a decreased probability that an avalanche will occur if the J_{inf} increases. From there, a random Gaussian distribution is created by using the inverse error function **(Press et. al., 2007)** by turning a random uniform distribution, with values between 0 and 1, into a Gaussian. The values in the Gaussian are h_i. Then the simulation initially sets all S_i = -1 and looks for the largest h_i in the Gaussian and increases h_i by H(t) to equal $H_{effective,i}$ so the first S_i flips starting an avalanche.

After the first S_i is flipped, each h_i in the Gaussian distribution is decreased by J_{inf} * q (q being the number of spins flipped). After all h_i are decreased, the code checks if any other S_i are at or above $H_{effective,i}$. If there is, then all $S_i > H_{effective,i}$ are set to +1 starting from largest to smallest. The code checks until no more S_i have flipped and then decreases the h_i by J_{inf} * q. Each time this process occurs is considered one time step for the avalanche duration. The total number of S_i flipped each time is the velocity for the time step. The avalanche continues with the loop until no S_i are at or above $H_{effective,i}$.

The h_i belonging to spins S_i that flipped in an avalanche are replaced by new (lower) h_i and the spins S_i = +1 that flipped up in the avalanche are replaced by new spins that point down S_i = -1. This way we model that the hosts that died in the most recent outbreak are replaced by newly

born living hosts with improved immunity. The h_i are replaced based on two methods (corresponding to different assumptions about the immunity of the descendants):

The first method (limited bounds method) randomly picks the new h_i from the h_i of the remaining living hosts left in the Gaussian (*limited bounds*). The idea is that the new descendants have approximately the same immunity as some of the still living hosts. To avoid having several hosts with identical immunity we pick a random h_i from the remaining living hosts and then set the new h_i to take a random value between the randomly picked h_i and the next largest h_i realized in the population of living hosts. The limited bounds method thereby does not allow for a new h_i to be smaller than the smallest h_i in the original Gaussian. This way the model assumes that the descendants have similar innate immunities as the rest of the living population.

The second method (unlimited bounds method) also creates new values of h_i to replace the old h_i values for the spins that flipped to $S_i = +1$. Again, the new h_i are drawn from the same Gaussian distribution (Press et. al., 2007**)** a**s** before. Again, we require that they cannot be larger than the largest remaining h_i of the remaining unflipped spins. However, in the unlimited bounds model we do allow the new h_i to be smaller than the smallest remaining h_i from our original Gaussian (i.e. thus the name *unlimited bounds model*). These requirements reflect the assumption that the descendants can be as immune or even more immune than the current living population. In this case, we expect the distribution of h_i to look more and more like the tail end of the Gaussian as time goes on.

In both versions of the model, after the h_i's of all spins that flipped in the previous avalanche are replaced, and the flipped spins are replaced with unflipped spins, the code starts a new avalanche by increasing H(t) just enough to flip only the spin with the largest h_i in the distribution of fields of unflipped spins (i.e. the least immune host gets infected). This event then

triggers the next outbreak and so on. The simulation repeats the process until the max time is reached, the avalanche's size is equal to the carrying capacity, or the avalanche size is greater than 10,000. At S > 10,000 the system has undergone a system spanning avalanche which was means to end the simulation. A system spanning avalanche is proportional to $\sqrt{N}$. For N = 316,228, the system spanning avalanches start around S = 500. The scaling is scene up to sizes of 10,000 so the simulation ends there.

Results

This mean field model has different regimes: for small infection rates (i.e. for small coupling J) or for large heterogeneity in the immunity (i.e. for a large width R of the Gaussian distribution of random fields $h_{i,t=0}$) or for large spacer acquisition rates (i.e. for large J_{inf}) the model predicts only small outbreaks that affect a microscopic fraction of the system. In the opposite limit, for high infection rates or for highly uniform immunity in the population, and for low spacer acquisition rates, it predicts macroscopically large outbreaks that wipe out a significant fraction of the population. In-between these two regimes there is a well-established critical regime where predictions are difficult because of an unusually high variability in the sizes of the outbreaks, which range from microscopic to macroscopic. The theory of phase transitions predicts that the scaling behavior of the outbreak statistics in this regime is universal, *i.e.* the same for many different diseases, and that it can be predicted with unusually simple models, as long as they capture the interaction range, symmetries, dimensions and a few other basic properties of the disease **(Sethna et., al. 2001).** The supplementary material shows a phase diagram that summarizes the trends in the outbreak statistics across the entire parameter range of the model. In the following, we mostly focus on the critical region and extract and test predictions for the outbreak statistics, which are marked by power law size and duration distributions for the outbreaks in this region. The model

predicts a universal set of scaling exponents τ, α, σνz etc. The exponents τ and α are the power law exponents of the complementary cumulative distribution functions (CCDFs) of avalanche sizes S and of avalanche durations T respectively. (An avalanche size is given by the number of lysed hosts in an outbreak). The complementary cumulative distribution function (CCDF) of the outbreak sizes S, called CCDF(S), gives the probability to get an outbreak larger than size S, and similarly the CCDF(T) gives the probability to have an outbreak that lasts longer than duration T. Within the scaling regime of sizes and durations we have CCDF(S) ~ $S^{(-\tau+1)}$ and CCDF(T) ~ $T^{(-\alpha+1)}$. The exponent σνz relates the avalanche size S to the avalanche duration T, as T ~ $S^{\sigma\nu z}$ see Eqns. (2.2-2.4) below and **(Sheikh et., al 2016)**. The predicted values of the scaling exponents are shown in **Table 2.1 (Alstott, et., al. 2014; Clauset et., al. 2009; Dahmen et., al. 1996)**. **Figure 2.1** shows statistical distributions from which the scaling exponents are derived. The power laws found in the limited bound (a) size vs duration, (c) size CCDF, and (d) duration CCDF plots of the avalanches reflect that for the used parameters the model is operating near an underlying critical point. The three fitted scaling exponents from the model simulations are listed in **Table 2.1.** They are close to the values obtained in mean field theory. In the scaling forms below f(x) and g(x) are (exponentially decaying) universal scaling functions, and S_{max} and T_{max} are the cutoffs to the power law distributions, *i.e.* the maximum avalanche size / avalanche duration respectively up to which the power law decay of the CCDF still holds to good approximation.

$$\boldsymbol{CCDF(S) = S^{(-\tau+1)} f(S/S_{max})} \quad \textbf{(2.2)}$$

$$\boldsymbol{CCDF(T) = T^{(-\alpha+1)} f(S/S_{max})} \quad \textbf{(2.3)}$$

$$\boldsymbol{T = S^{\sigma\nu z}} \quad \textbf{(2.4)}$$

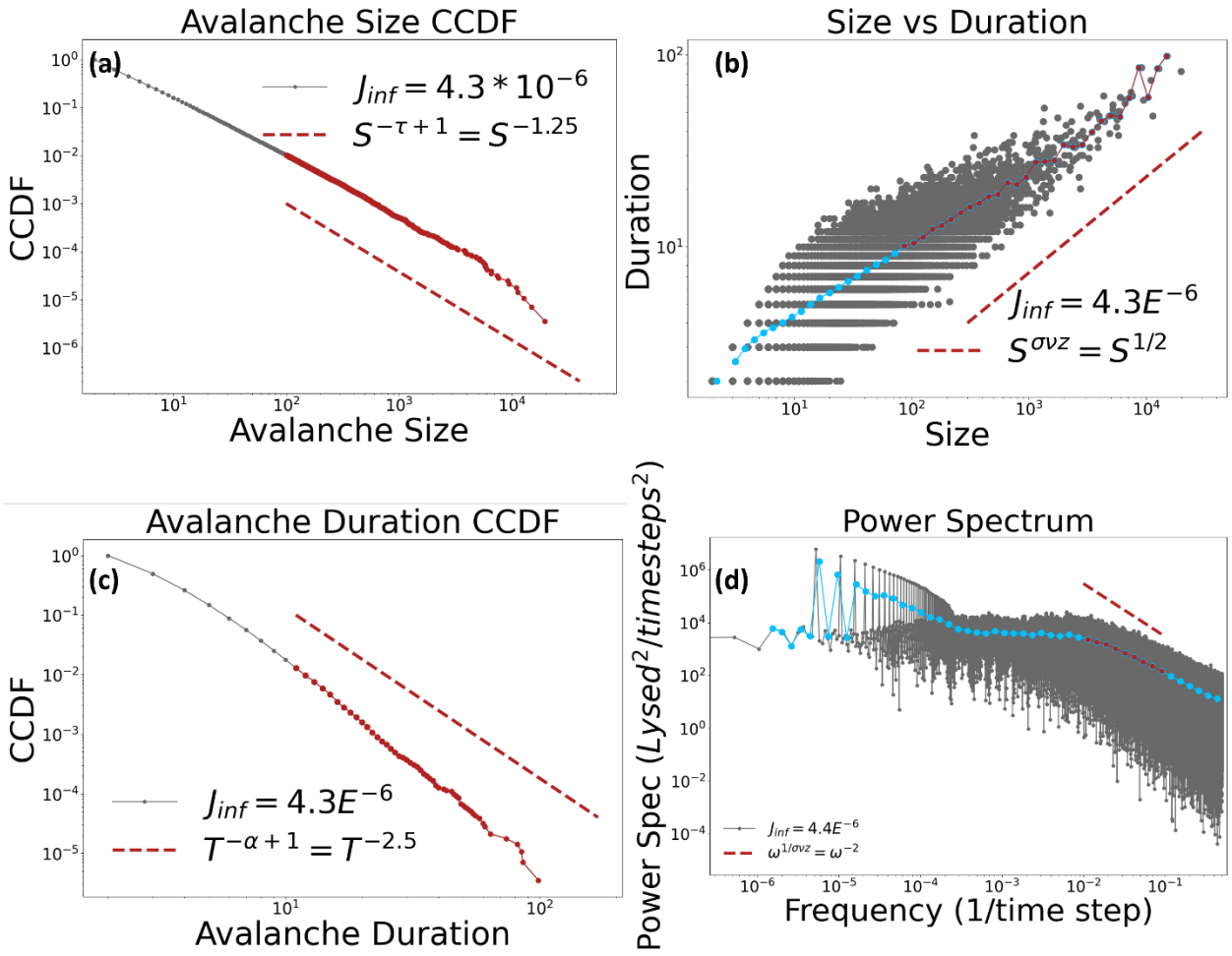


Figure 2.1: Limited bounds simulation: Scaling regimes from the (a) Size CCDF, (b) Duration versus size (number of bins = 50), (c) duration CCDF, and (d) power spectrum of the limited bounds simulation. The red line indicates the analytically predicted mean field power law **(Dahmen, et al., 1996; Dahmen et., al. 2016; Sethna, et., al 2001).** The exponents σνz, τ, α give the power laws of the size versus duration, size CCDF, and duration CCDF, respectively. The power spectra are the absolute square of the Fourier transform of the time trace of the new infections per time step. The system size was N =316,228 and the system max time was 300,000 timesteps; J_{inf} = 4.3 * 10^{-6} and J =1 and R = 1.12.

The self-similarity of the avalanches is further demonstrated by the scaling collapses: (1) for a collapse of the avalanche size distributions at different values of J_{inf} (see Figure 2, where $\langle S^2 \rangle$ is the mean square avalanche size), and (2) for scaling collapses of the average temporal avalanche profiles ("shapes") at fixed size ($\langle V(t|S) \rangle$) and at fixed duration ($\langle V(t|T) \rangle$) (see Figure 3). The

average temporal profiles at fixed size and duration are obtained by averaging the time traces of the number of new infections per time step V(t) during an outbreak for all avalanches of the same size S or duration T respectively. Equations 2.5-2.8 give the scaling forms for the average time traces <V(t|S)> and <V(t|T)> respectively. At the critical point, the average temporal profiles at different sizes or at different durations collapse onto one another, proving self-similarity of the dynamics on different scales. Not only the scaling exponents but also the collapse functions are predicted to be universal, *i.e.* independent of the microscopic details. The exponents used for the scaling collapses are given in the Figures and included in Table 1. The exponents in **Table 2.1** reflect the best fit exponents obtained from the CCDFs. They agree with the exponents obtained from the collapses within the error bars. We use the known scaling forms $\langle V(t|S)\rangle \sim S^{1-\sigma\nu z} F_S(t/S^{\sigma\nu z})$ and $\langle V(t|T)\rangle \sim T^{1/\sigma\nu z-1} F_T(t/T)$ where $F_s(x)\sim x \exp(-x^2/B)$ with B a nonuniversal constant and $F_T(x)\sim x(1-x)$ in mean field theory are universal scaling functions (Sethna et. al., 2001). We use following rescalings to obtain the corresponding collapses of the average outbreak growth rates onto the universal scaling functions: For the average outbreak growth rate <V(t,S)>:

rescaled x-axis = $t/S^{\sigma\nu z}$ **(2.5)**

rescaled y-axis = $\langle V(t|S)\rangle/S^{1-\sigma\nu z}$ **(2.6)**

For the average outbreak growth rate <V(t,T)>:

rescaled x-axis = t/T **(2.7)**

rescaled y-axis = $\langle V(t|T)\rangle/T^{1/\sigma\nu z - 1}$ **(2.8)**

As expected, similar behavior is observed in the unlimited bounds simulation. **Figure 2.4** shows the (a) size vs duration, (c) size CCDF, and (d) duration CCDF graphs for the unlimited bounds simulations. Here, for N=316,228 spins the apparent critical point for $J \equiv 1$ is at $J_{inf} = 5.2$

* 10^{-6}, and disorder R = 1.2 compared to the J_{inf} of the limited case (4.3 * 10^{-6}) with disorder R = 1.2. The simulation maximum time was set higher because it takes more timesteps before larger avalanches occur due to the different method of field replacement. The scaling exponents for σνz and τ are similar in both cases, however α appears larger in the limited case compared to the unlimited case. Note, however, that the scaling regime for the duration distribution is small and the error bars in both cases are overlapping with each other. Theoretically, based on renormalization group arguments (**Sethna et. al., 2001**), in the thermodynamic limit we expect the exponents in both cases to be the same. **Figure 2.5** depicts a collapse of size CCDFs at different J_{inf} larger than 5.2 * 10^{-6}. The scaling form for the CCDFs can be written as CCDF(S) ~ $< S^2 >^{\frac{1-\tau}{3-\tau}} F_{CCDF}(\frac{S}{<S^2>^{\frac{1}{3-\tau}}})$ with universal scaling function F_{CCDF} **(Uhl et al. 1015).** The curves fall on top of one another when the axes are rescaled according to Equations 2.9 and 2.10 **(Uhl et. al., 2015).**

rescaled size = $\frac{S}{<S^2>^{\frac{1}{3-\tau}}}$ **(2.9)**

rescaled CCDF(S) = $\frac{CCDF(S)}{<S^2>^{\frac{1-\tau}{3-\tau}}}$ **(2.10)**

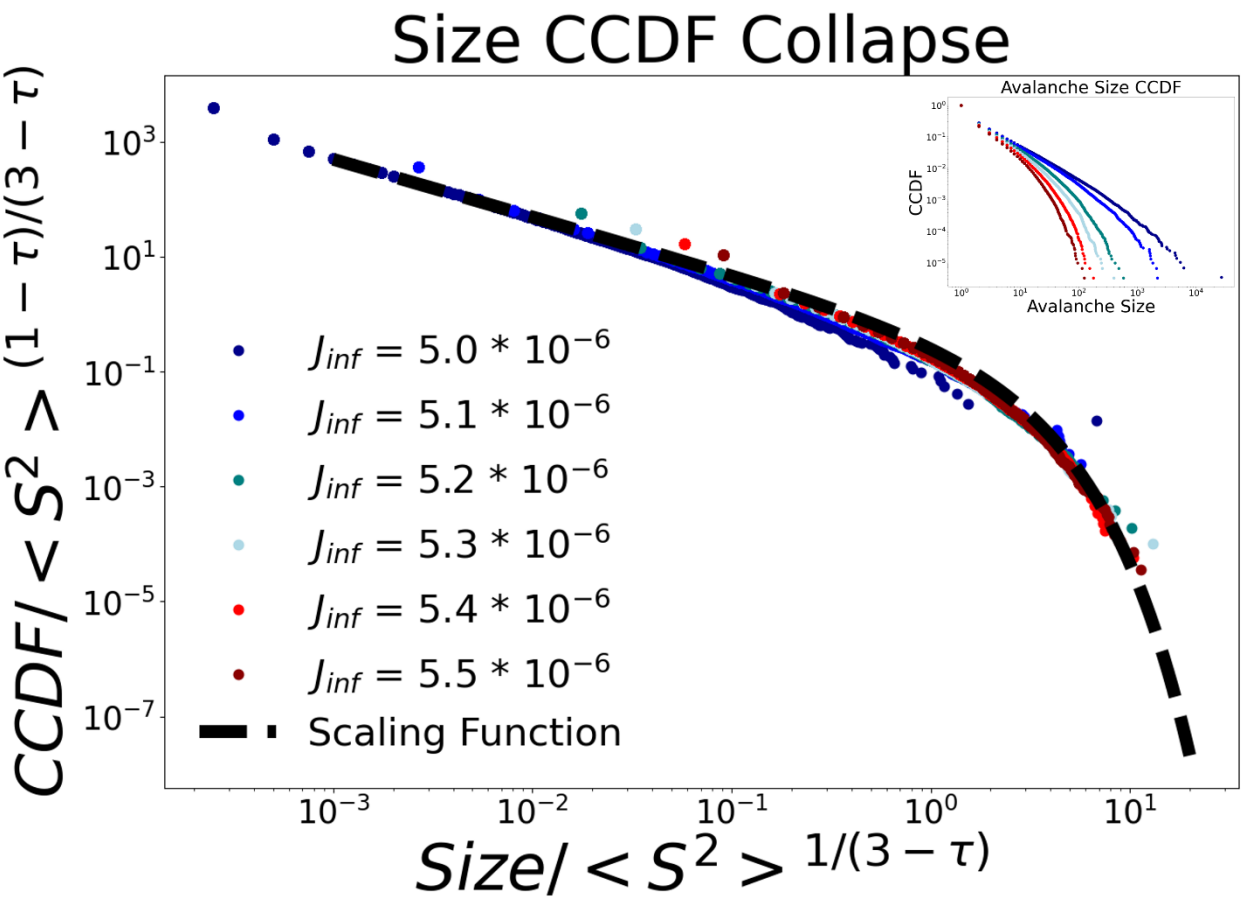


Figure 2.2: Main Figure: Limited bounds simulation: Scaling collapse of the CCDF of avalanche sizes for various J_{inf} above criticality of the limited simulation. For the collapse $\tau = 2 \pm \frac{.25}{.25}$ was used. The inset shows the unscaled CCDFs for different values of J_{inf} given in the legend. The system size was N = 316,228 spins with disorder (R) = 1.12. The nearest neighbor coupling was set to J=1. The maximum simulation time was set to 300,000-time steps. The form of the fitted scaling function is $\frac{e^{(-Ax)}}{Bx}$ with A = .5 and B = 3 (dashed line).

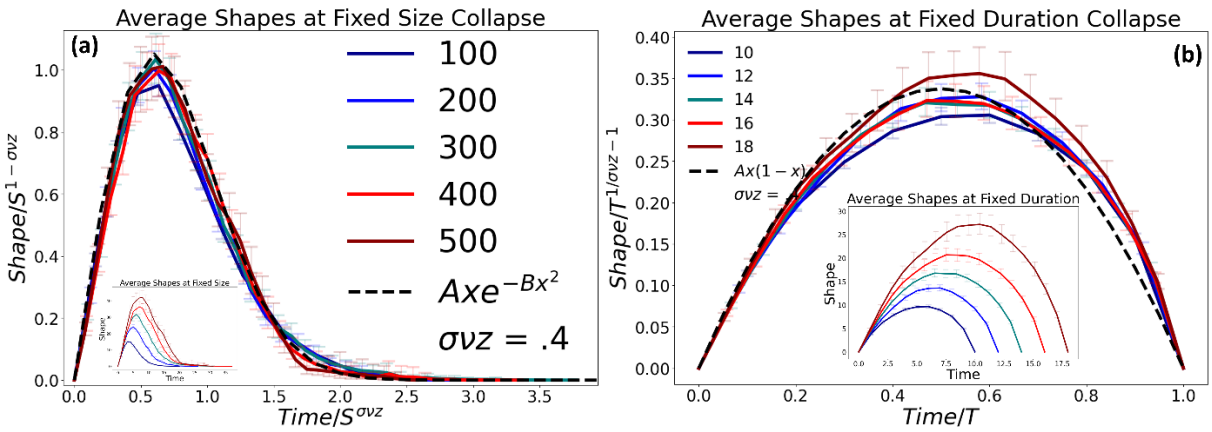


Figure 2.3: Limited bounds simulation: The main panels show collapses of the avalanches shapes at (a) fixed sizes and (b) fixed durations. The insets show the unscaled shapes at different sizes (a) or different durations (b). The system size was N=316,228 with disorder (R) = 1.12, and J = 1. The maximum simulation time was set to 300,000 time steps. The black dashed line in (a) shows the predicted mean field scaling collapse function ($Axe^{-(Bx^2)}$(A = 2.5 and B = 1.3)). The black dashed line in (b) shows the predicted mean field scaling collapse function (Ax(1−x) plotted (A = 1)). The average shapes are from 10 simulation runs with J_{inf} = 4.3 * 10^{-6}. A value of σνz = .38 ± $\frac{.02}{.03}$ was used to collapse the shapes at fixed size and fixed duration curves, respectively.

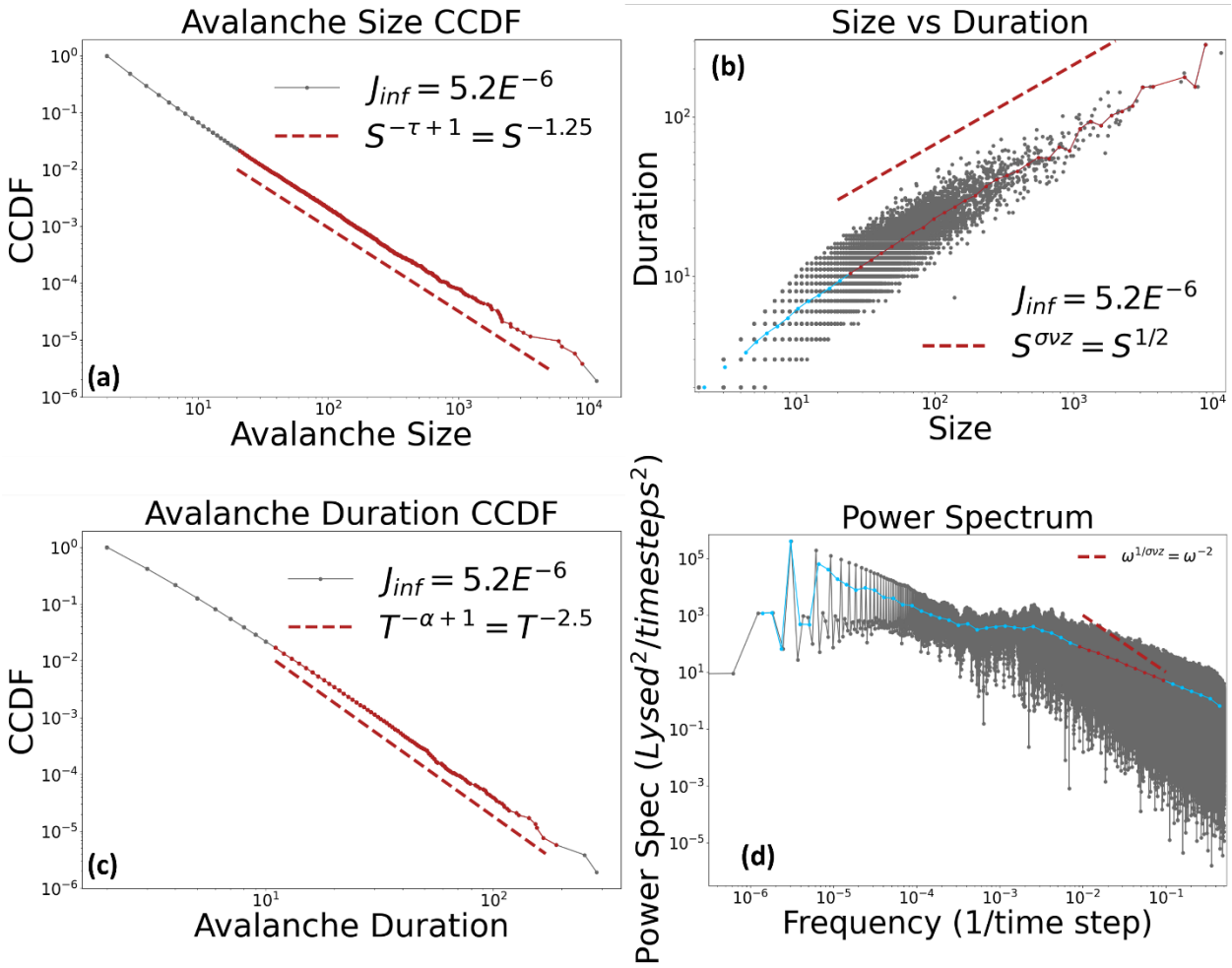


Figure 2.4: Unlimited bounds simulation: Scaling regimes from the (a) Size CCDF, b) size versus duration, and (c) duration CCDF (d) power spectrum. The red line indicates the analytically predicted mean field power law. Exponents σνz, τ, and α are derived from the power laws of the size vs duration, size CCDF, and duration CCDF, respectively. The system size was N=316,228 spins with disorder (R) = 1.12, J=1, J_{inf} = 5.2 * 10^{-6}. The maximum time was set to 900,000 time steps. Exponents are

from data **Figure 2.4 (cont.);** across 5 runs with different realizations of the disorder.

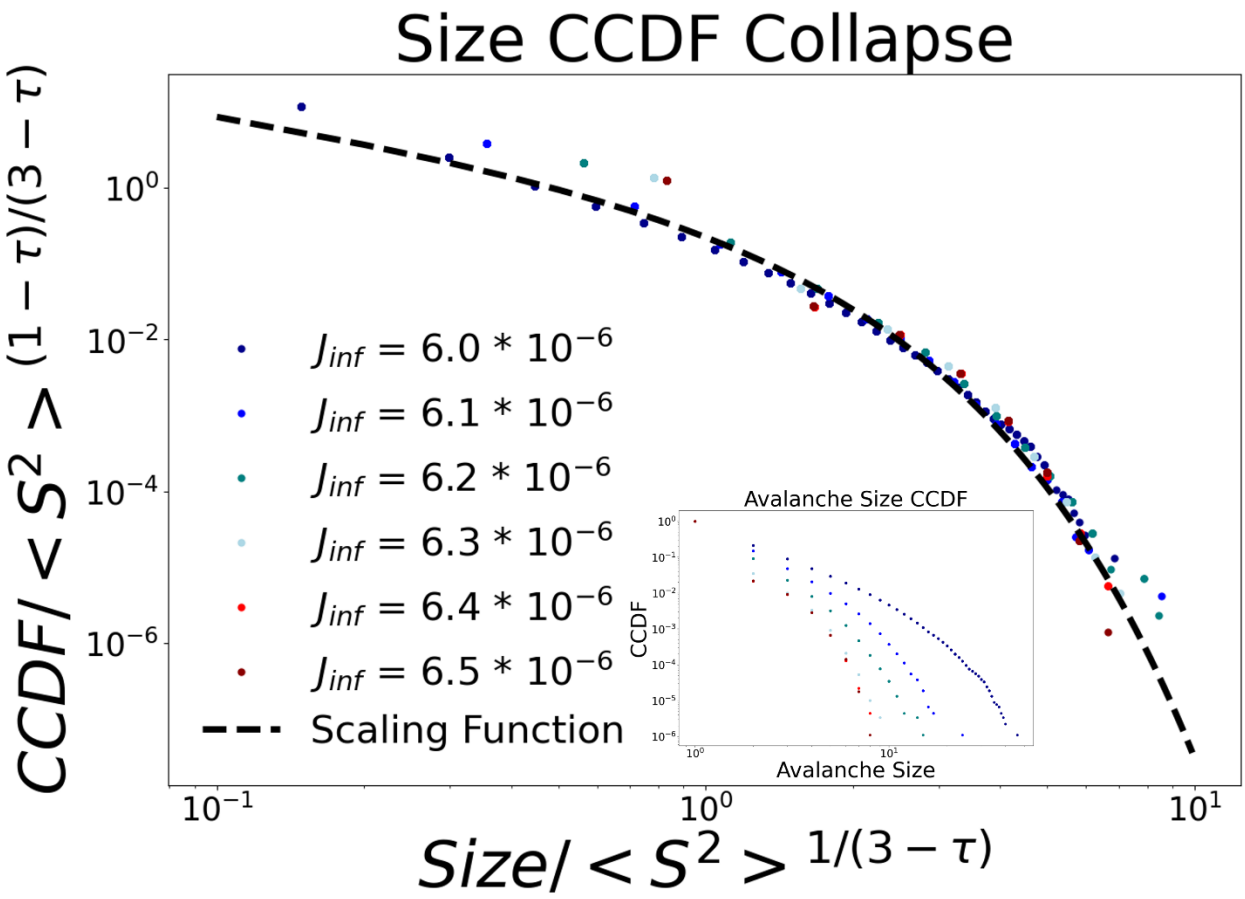


Figure 2.5: Unlimited bounds simulation: The CCDF of avalanche sizes for various Spacer acquisition probabilities (SAPs) J_{inf} above criticality. We used $\tau = 2.2 \pm \frac{.15}{.15}$ for the collapse. The inset shows the same CCDF with the x and y axis scaled according to the predicted scaling equation **(Equs. 13, 14)**. The system size was N = 316,228 spins with disorder (R) = 1.12, and J=1. The maximum time was set to 900,000-time steps. Exponents are from data across 10 runs with $J_{inf} = 5.2 * 10^{-6}$. The fitted scaling function is $\frac{e^{(-Ax)}}{Bx}$ with A = .5 and B = 3 (dashed line).

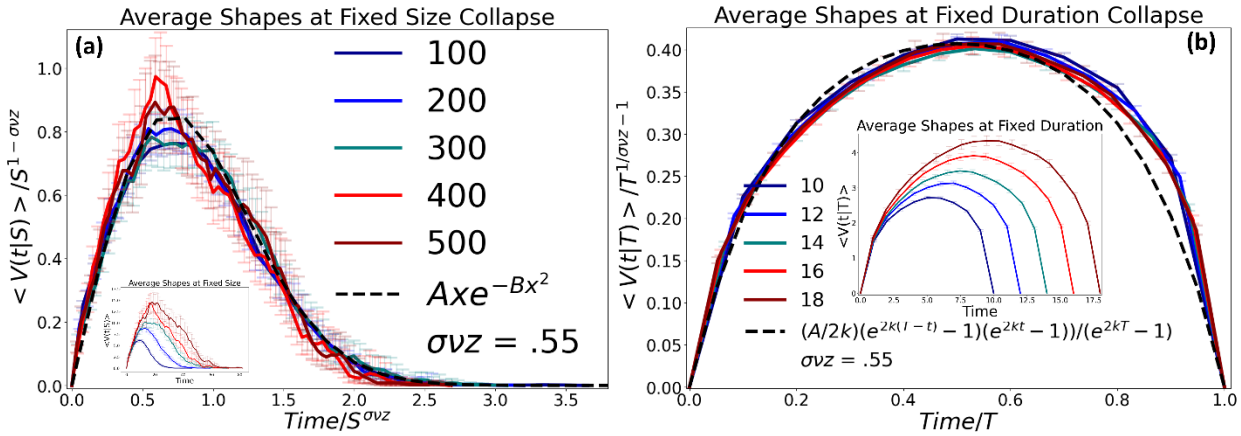


Figure 2.6: Unlimited bounds simulation: The main panels show collapses of the average temporal avalanche profiles,

Figure 2.6 (cont.): averaged at (a) fixed sizes and (b) fixed durations of the avalanches, with different sizes (a) and different durations (b) corresponding to each curve. Insets show the unscaled shapes at different sizes or durations, respectively. The system size was N=316,228 spins with disorder (R) = 1.12 and J=1. The maximum time was set to 900,000-time steps. The black dashed line in (a) shows the predicted mean field scaling collapse function ($Axe^{(-Bx^2)}$ (A = 3.8 and B = 2)**[4]**). The black dashed line in (b) shows the predicted mean field scaling collapse function ($\frac{A}{2k}\frac{e^{2k(T-t)-1}(e^{2kt}-1)}{e^{2kT}-1}$ $(Uhl\ et.al.,2015)$ plotted (A = 4, k = 3, T = 1)). Exponents are from data across 10 runs with J_{inf} = 5.2 * 10^{-6}. The value of σνz =.60 $\pm\frac{.02}{.07}$ was used for both scaling collapses.

Table 2.1	$\boldsymbol{\tau}$	$\boldsymbol{\alpha}$	$\boldsymbol{\sigma\nu z}$	$\boldsymbol{\frac{\tau-1}{\alpha-1}=\sigma\nu z}$
Limited	$\mathbf{2.20 \pm \frac{.12}{.08}}$	$\mathbf{5.02 \pm \frac{.15}{.15}}$	$\mathbf{.36 \pm \frac{.03}{.04}}$	$\mathbf{.30 \pm \frac{.03}{.02}}$
Unlimited	$\mathbf{2.42 \pm \frac{.15}{.23}}$	$\mathbf{3.95 \pm \frac{.12}{.10}}$	$\mathbf{.52 \pm \frac{.02}{.10}}$	$\mathbf{.48 \pm \frac{.05}{.08}}$
Mean Field Integrated	**2.25**	**3**	**.5**	**.5**

Table 2.1: Theoretical and measured scaling exponents: Listed theoretical and measured values for the scaling exponents τ, α, and σνz for the simulations.

The model was run with various system sizes to show that scaling is present for various populations sizes. **Figure 2.1** shows that systems sizes 2x and 4x the original show a powerlaw in the size CCDFs for both the limited and unlimited methods. Note that J_{inf} gets smaller as the system size increases. A phase diagram (**Figure 2.2 a,b**) shows the span of criticality across various Rs and s for the limited and unlimited methods, respectively. The phase diagrams show that R and are inversely proportional.

Discussion

The scaling regimes span multiple decades for the size vs duration, size, and duration distributions indicating the self-similarity of the avalanches extends for a fairly broad range of scales. The average shapes at fixed size and fixed durations collapse onto one another for avalanches within the scaling regime as is shown by appropriate scaling collapses**.** The scaling

collapses of the temporal avalanche shapes use the value for the exponent $\sigma\nu z$ that was derived from the size vs duration power law. The collapse of the size CCDFs with J_{inf} greater than the critical value $J_{inf,critical}$ (limited model $J_{inf,critical} = 4.3 * 10^{-6}$; unlimited model $J_{inf,critical} = 5.2 * 10^{-6}$) is based on the mean field scaling **equations (2.9,2.10)**. The results imply that the mean field model can be used to understand the scaling aspects of the statistics and dynamics of the outbreaks. Both methods of host immunity replacement (limited model and unlimited model) show these results at similar J_{inf}. Whether the limited bounds or unlimited bounds case describes the proper host population dynamics depends on the details of how immunity is passed on to the next generation – either way we expect the same scaling behavior in both cases.

The model was run with various system sizes to show that scaling is present for various populations sizes. **Figure 2.1** shows that systems sizes 2x and 4x the original show a powerlaw in the size CCDFs for both the limited and unlimited methods. Note that J_{inf} at the critical point gets smaller as the system size increases, as is to be expected (Kuntz et. al., 2001). A phase diagram (**Figure 2.3 a,b**) shows the span of criticality across various disorder values R and J_{inf} for the limited and unlimited respectively. The phase diagrams show that for the phase boundary $R \sim J - J_{inf}$, as is to be expected from mean field theory. The red dot on each phase diagram is the parameter combination of R and J_{inf} used to find critical scaling in the simulation for J = 1. The CCDF for J_{inf} above and below criticality are plotted within the respective areas of the phase diagram. Figure 2.3 shows the average shapes profile at fixed avalanche sizes for avalanches that are greater than the scaling regime. The avalanches show a Gaussian shape compared to the avalanches within the scaling regime which are more fluctuating individually and have a different average shape. The red circle in the insert denotes the avalanche sizes that were used for the shapes profiles.

Conclusion

It is important to have efficient resource allocation because an endemic can happen at any time and predictions for the outbreak statistics would be needed for resource allocations. Once a disease's parameters are found, the statistics and the dynamics of even the largest outbreaks of an endemic can be calculated. As shown in detail in Figure 2.3 the mean field model provides tools to deduce if the initial infection rates will propagate into a large outbreak within the scaling regime or grow even larger into a catastrophically large event. This is possible because the model predicts that the temporal profiles of the catastrophically large events are bell-shaped with statistical properties that differ from the strongly fluctuating profiles of the smaller outbreaks in the scaling regime. Furthermore, the outbreak size distributions and duration distributions of a novel, infectious disease can be calculated if the parameters (infection rate and re-infection) have been measured. Such information is important for preparing sufficient hospital capacities etc. ahead of an outbreak. Future work will refine the models to increase the accuracy of the predictions. **Figure 2.9** shows Gaussian shape profiles <V(t)> for the large outbreaks for model parameters away from criticality, in the large event regime. These shapes clearly differ from the shapes seen in **Figure 2.3.** The model thus shows that the early dynamics of the outbreaks can be used to predict if a starting outbreak could turn into a catastrophically large event or not.

Lastly, this model could be used to study the effect of countermeasures. For example, if people were asked to stay home and isolate once there is a perceived risk of a large outbreak, the model dynamics may be switching to the two-dimensional case with only nearest neighbor interactions. The two-dimensional model is slightly different from the presented well mixed case. The 2D model exponents have been computed previously for 2 dimensional ferromagnets (Kuntz et. al., 2001), and will be adjusted for modeling diseases in future studies.

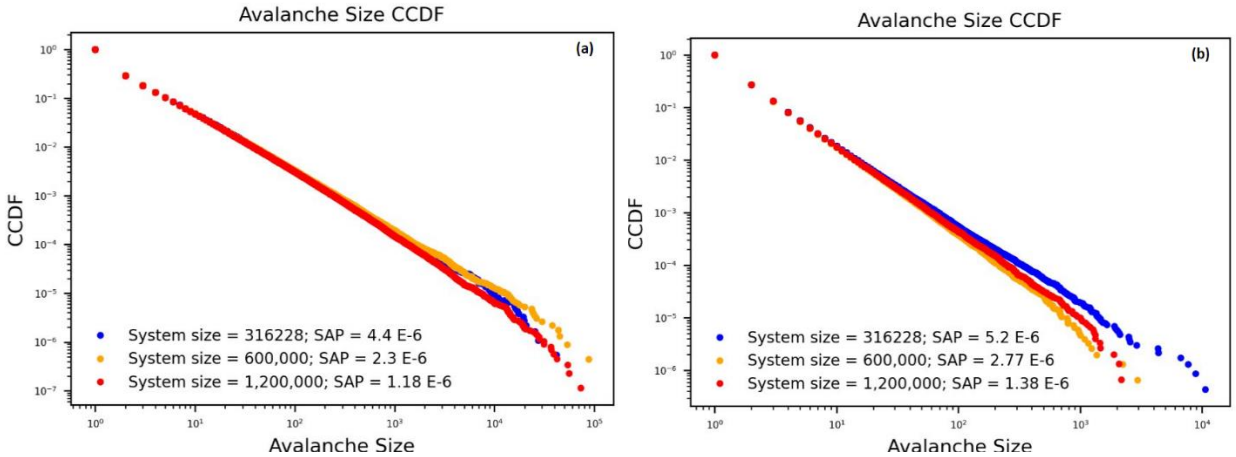


Figure 2.7: The size CCDF at various system sizes with J=1 and R = 1.12 for the (a) limited method $J_{inf} = 4.4E^{-6}$ had a system size of 316,228 and max time of 250,000 time steps. $J_{inf} = 2.3E^{-6}$ had a system size of 600,000 and a max time of 400,000 time steps. $J_{inf} = 1.2E^{-6}$ had a system size of 1,200,000 and a max time of 1,000,000 time steps. (b) is the unlimited method. $J_{inf} = 5.2E^{-6}$ had a system size of 316,228 and max time of 1,000,000 time steps. $J_{inf} = 2.8E^{-6}$ had a system size of 600,000 and a max time of 2,000,000 time steps. $J_{inf} = 1.4E^{-4}$ had a system size of 1,200,000 and a max time of 2,000,000.

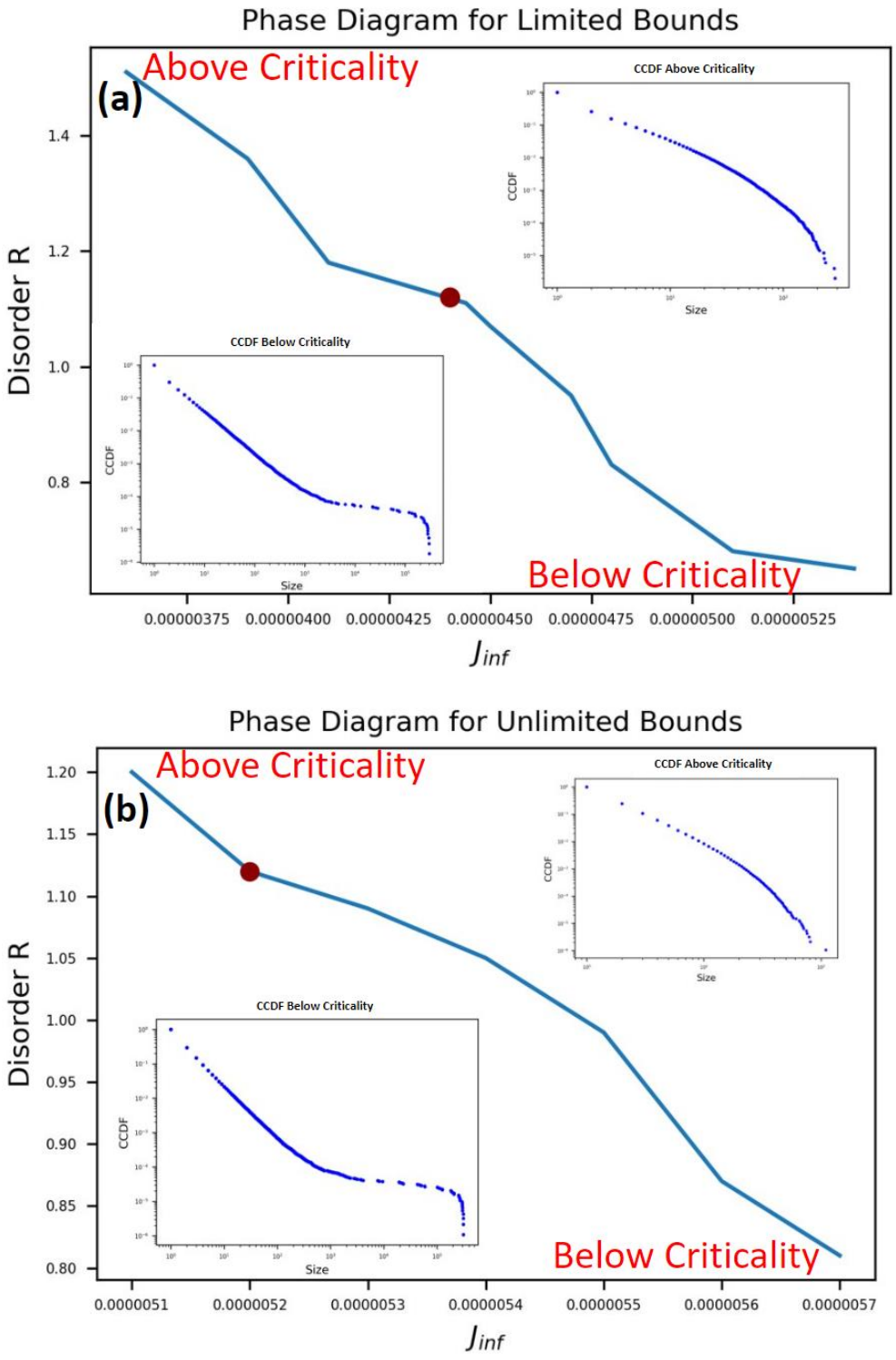


Figure 2.8: The parameter combinations of R and Jinf that result in criticality for the (a) limited and (b) unlimited methods. The region labeled “Above Criticality” (top region) represents parameter combinations that result only in smaller events and the region named “Below Criticality” (lower region) refers to a regime where the system also exhibits macroscopically large runaway

Figure 2.8 (cont); avalanches. The red dot shows the critical point where power law scaling is observed. Insets are CCDFs with Jinf above and below Jinf,critical, $J = 1$, and the system size for each simulation was 316,228.

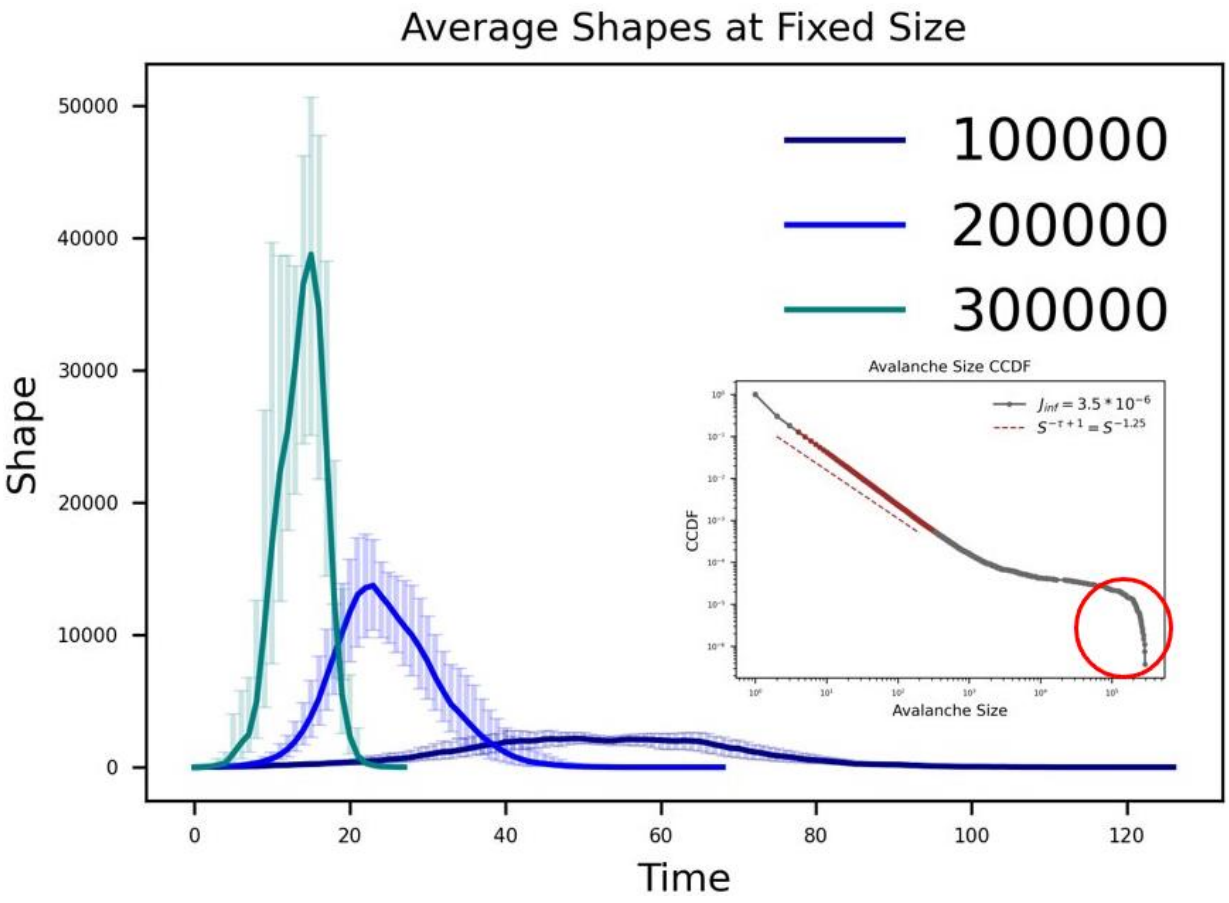


Figure 2.9: Temporal profile of avalanches larger than those in the scaling regime: (avalanches shapes) for fixed avalanche sizes for the limited bounds model above the scaling regime. The inset is the Size CCDF of a system below criticality, where the system shows macroscopic runaway avalanches. The red circle is the region of sizes used for the average shapes profiles of the runaway events that are shown in the main figure. The system size was N=316,228 with disorder (R) = 1.12, J=1, $J_{inf} = 3.5E^{-6}$. The maximum time was set to 500,000 time steps over 20 runs.

Chapter 3: Fitting Lotka-Volterra host population dynamics with the mean field

Introduction

Antibiotic resistance is an ever-looming danger in our lives. Bacteria are mutating and becoming resistant to antibiotic treatments at a fast rate. Phage therapy is thought to be an alternative solution to antibiotic treatments of bacterial infections (d'Herelle et. Al., 1926, Thacker et. al., 2003). This therapy is the idea that viral phages can mutate past bacteria immunity thus killing the bacterial infection because phages have exponential population growth dynamics (Schoolnik et. al., 2004, Kysela et. al., 2007). Similarly, microbial diversification is crucial for viral resistance in a population. One method in viral immunity is for microbials to use CRISPR and the CRISPR proteins (Cas-9) to obtain immunity from viruses (Labrie et al., 2010, van Houte et al., 2016, Mayer et al., 2016). The CRISPR-Cas systems can acquire small fragments of nucleic acids, "protospacers", from infectious viral RNA and incorporate them into its own genomic code as "spacers" resulting in the host gaining immunity from said viral infection. The spacer is passed down to future generations along with viral immunity (McGinn et., al. 2019). Bacteria are immune to the virus if they have a spacer that matches an infecting viruses' protospacers.

These two dynamics control how the viral and host abundances behave in a community. Pascal et al. (2020) have modeled these dynamics with Lotka-Volterra differential equations. Their simulations show areas where the virus outperforms the host causing drastic drops in the host population and areas where the host suppresses the virus causing the hosts to remain near their carrying capacity. The arms race between bacteria immunity and viral mutations in these two areas has motivated us to look at the dynamics of host population growth and decay. The model discussed here indicates that population growth and decay is reminiscent of avalanches. Avalanches are intermitted, fast changes to a system that is experiencing a slowly increasing

driving force (Sheikh et. al, 2016). There are many systems that produce avalanches or crackling noise. Several power law distributions are used to determine the characteristics of crackling noise that can help identify the dynamics of the noise (Uhl et al., 2015 and references therein).

In this chapter, we show how the outbreaks in the host dynamics reflect avalanche dynamics. We first show, using Pascal's et al.'s model, that host population bursts can be categorized as avalanches. Next, we show that the system has scaling regimes according to crackling noise theory, and that host population dynamics show self-similarity of the small and large avalanches in the scaling regime. This work provides a new understanding of how bacteria and virus diversification can play out in an ever-changing ecosystem. The virus host relationship is important for the world's ecosystem, because the evolution, transmission dynamics, and population stability of each is affected by the arms race between the virus and the host.

Results:

Avalanche statistics often show power law scaling within a certain scaling regime of the experimental parameters. The scaling regime is the regime of sizes and durations where avalanches are self-similar on a range of different scales. For example, power law scaling is seen in scatter plots of the avalanche size vs its duration, in complementary distribution function (CDF) of the size, the CDF of durations, and the power spectrum of the time trace of the avalanche propagation speed. In the scaling regime many statistical properties show power law scaling with universal critical exponents that are expected to be the same for many different systems (Uhl et. al. 2015). **Figure 3.1** demonstrates two distinct regimes in the size CCDF (3.1a) and duration vs size (3.1b) resulting in the naming of the "small regime" and "large regime." The small regime is the first powerlaw in the size CCDF and any avalanche $S < 3{,}000$ whereas the large regime is the second powerlaw in the size CCDF and any avalanche $S > 3{,}000$.

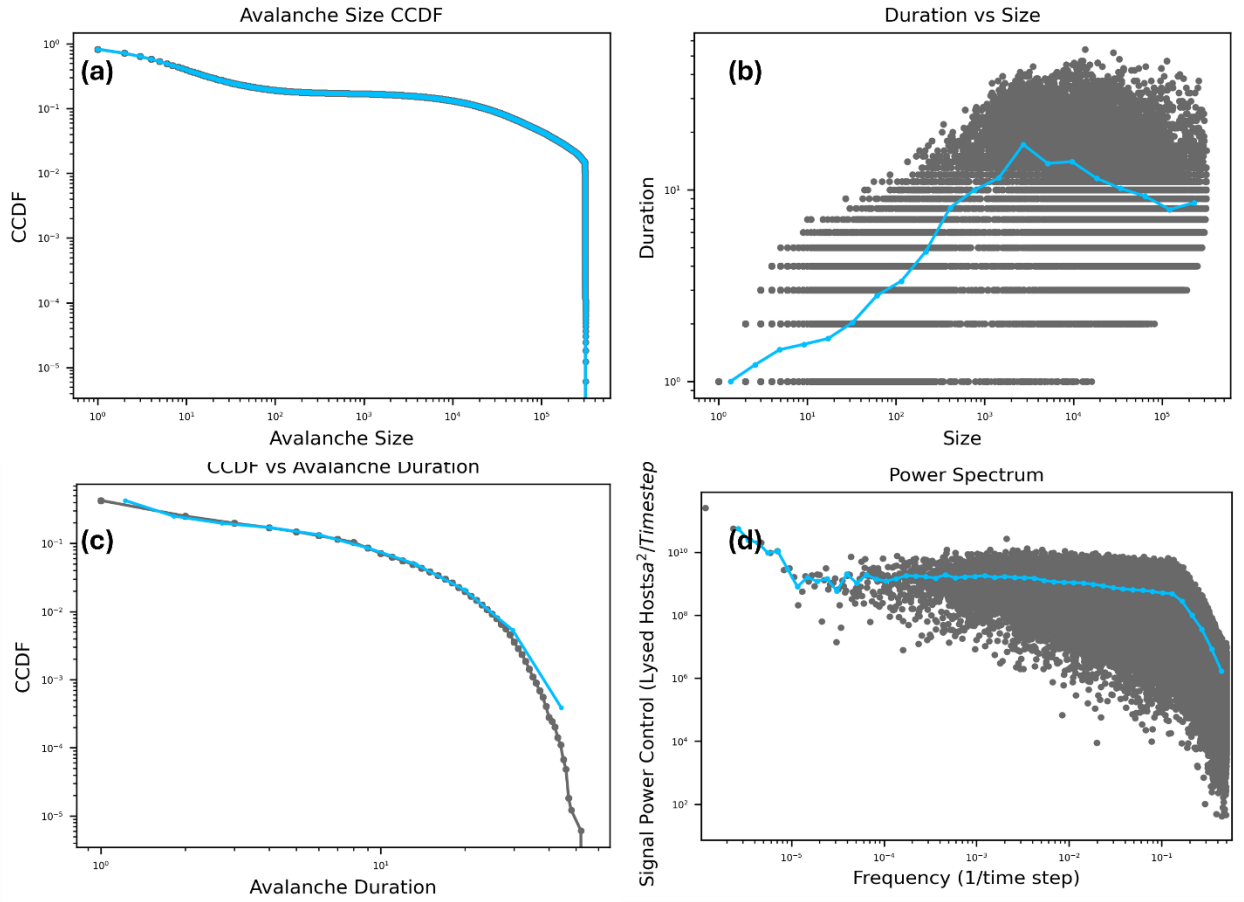


Figure 3.1: Scaling regimes of host burst dynamics. The common graphs for identifying scaling regimes for all host avalanches within a system. a) CCDF of host avalanche sizes, b) the host avalanche duration vs. size, c) host avalanche duration CCDF, and d) the power spectrum of the power spectra is the absolute square of the Fourier transform of the time trace of the newly lysed hosts per time step. The virus mutation rate is $1.17E^{-6}$. The duration vs size shows a change in sign for the correlation between duration and size resulting into two different regimes around S = 5000. The system size was N hosts = 316,228.

Our data shows that power laws describe the statistical distributions associated with the host population within a broad range of avalanche sizes and durations. The CCDF of avalanche sizes (**fig. 3.2a**) has a scaling exponent around $1.27 \pm \frac{.34}{.14}$ and spans across 1.5 decades. The avalanche duration vs avalanche size has a scaling exponent around $.43 \pm \frac{.07}{.17}$ roughly 1.5 decades

(**fig. 3.2b**). The CCDF of the avalanche durations has a scaling exponent around 1.97 $\pm\frac{.35}{.97}$ and spans 1 decade (**fig. 3.2c**). Lastly, the power spectrum of the number of new hosts lysed versus time also shows a power law. (**fig. 3.2d**). The power law exponents are important because these scaling exponents reveal the underlying universality class (described in chapter 2). They are independent of the microscopic details of the system and should therefore be the same for many different diseases. Also, they reveal the degree of self-similarity that is present across a certain range of avalanche sizes, and they are useful for estimating the risks of large outbreaks from the statistics of the much more abundant small outbreaks. The scaling exponents $\sigma\nu z$, **τ**, and **α** are derived from the power laws of the size vs duration plot, the CCDF of sizes, and the CCDF of durations, respectively. The exponent values are listed in **table 3.1**.

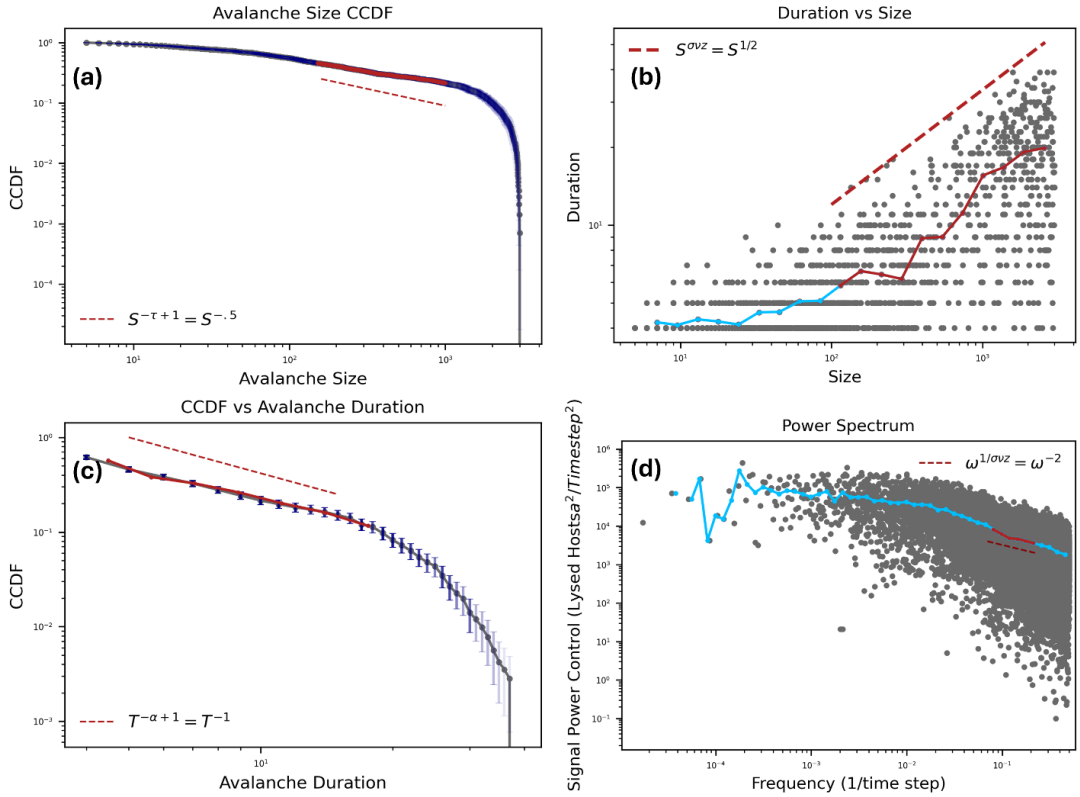


Figure 3.2: Small regimes of host burst dynamics. The statistics demonstrate a scaling regime (red line) across the loglog plot of the a) CCDF of host avalanche sizes, b) the host avalanche duration vs. size, c) host avalanche duration CCDF, and d) the power

Figure 3.2 (cont.); spectrum of the power spectra is the absolute square of the Fourier transform of the time trace of the new lysed hosts per time step. A virus mutation rate of $1.17E^{-6}$. The exponents σvz, τ, and α give the power laws of the size versus duration, size CCDF, and duration CCDF, respectively. The power spectra are the absolute square of the Fourier transform of the time trace of the new infections per time step. The system size was N hosts = 316,228.

Exponents	$\boldsymbol{\tau}$ **(size dist.)**	$\boldsymbol{\alpha}$ **(duration dist.)**	$\boldsymbol{\sigma\nu z}$ **(size vs duration)**	$\boldsymbol{\frac{\tau - 1}{\alpha - 1} = \sigma\nu z}$
Host Avalanches	$\mathbf{1.27 \pm \frac{.34}{.14}}$	$\mathbf{1.97 \pm \frac{.35}{.97}}$	$\mathbf{.43 \pm \frac{.07}{.17}}$	$\mathbf{.282 \pm \frac{.27}{.08}}$
Mean Field Theory	**1.5**	**2**	**.5**	**.5**

Table 3.1: Scaling exponent values. This table lists the values of each scaling exponent for the host avalanches. The mean field exponents are listed for comparison. These values were obtained using the slopes of the graphs in Figure 2 which use all spatial acquisition probabilities and a virus mutation rate of $1.17E^{-6}$. I used bootstrapping to calculate the exponents and error bars (Alstott et. al., 2014; Clauset et. at., 2009)

Self-similarity is reflected by power law scaling. Here, self-similarity means that avalanches across various sizes show similar dynamics. Power law distributions alone are often not sufficient to show self-similarity. The self-similarity can be shown in more detail if CCDFs and the average temporal profiles ("shapes") of the avalanches size and durations can be collapsed upon one another. The avalanches will collapse according to crackling noise theory if the critical exponents of the theory are used to scale the *x* and *y*-axis of the plots mentioned to get a scaling collapse. The three exponents used are **τ**, **α**, and $\sigma\nu z$ calculated from the size CCDF, duration CCDF, and the duration versus size, respectively. From the slope of the CCDF, **τ** and **α** are extracted from the plots according to the scaling laws of Equs. 2.2 and 2.3 for their respected CCDFs. The scaling exponent $\sigma\nu z$ is the slope of the log log plot of duration vs. size as given in Equ. 2.4. For avalanches $\sigma\nu z$ usually is the inverse of the negative slope of the log log plot of the

power spectrum. Furthermore, for avalanches $\frac{\boldsymbol{\tau-1}}{\boldsymbol{\alpha-1}}$ should equal $\sigma\nu z$ (Coleman et. al., 2017). Mean field theory exponents $\boldsymbol{\tau}$, $\boldsymbol{\alpha}$, $\sigma\nu z$, have values of 1.5, 2, and .5 (see T**able 3.1**). The mean field exponents are consistent in value to the ones calculated from the host abundance data, indicating that the data does follow mean field avalanche theory. This conclusion is strongly supported also by scaling functions from the dynamics that also agree with the mean field theory predictions:

The average temporal profiles ("shapes") at fixed avalanche duration and at fixed avalanche size respectively are plotted against time to show universal aspects of the avalanche dynamics. The size and duration CCDF scaling regime were binned into three bins equally spaced (200, 305, 500, 740, 900) in size and (8, 10, 12, 14, 16) in duration, respectively. Then they were plotted to show how the shapes of the small avalanches differ from those of the larger avalanches in the scaling regime. More importantly, criticality requires that these shapes collapse onto one another if the axes of the plots are rescaled according to the mean field scaling results. The *x* and *y*-axis of the average shape with fixed size (S) are scaled, with time (t), by $t/S^{\boldsymbol{\sigma\nu z}}$ and $shape/S^{1-\boldsymbol{\sigma\nu z}}$, respectively. Similarly, the *x* and *y*-axis of the shapes at fixed duration (T) are scaled by t/T and $shape/T^{(1/\boldsymbol{\sigma\nu z})-1}$. After the axis is scaled in this way, the curves collapse onto one another and onto the predicted mean field scaling functions.

Figure 3.3a and 3.3b show shape collapses at fixed duration and fixed size, respectively. The inserts of each figure are the raw shapes before the collapse (i.e. before rescaling the axes). The exponent $\sigma\nu z$ is used to scale the y-axis by $T^{(1/\sigma\nu z)-1}$ and to collapse the shapes for fixed duration. The resulting value for $(1/\sigma\nu z)-1$ is close to the mean field value of 1. For the shapes at fixed size the exponent $\sigma\nu z$ used for the rescaling of the *x* and *y* axis' is $\sigma\nu z = 0.65 \pm \frac{.45}{.15}$. The

value is close, with error bars, to the measured value of $.43 \pm \frac{.07}{.17}$.. The measured σνz is close to the mean field value of 0.5. The quality of the shape collapse supports the hypothesis that there is self-similarity in the avalanche sizes across the scaling regime.

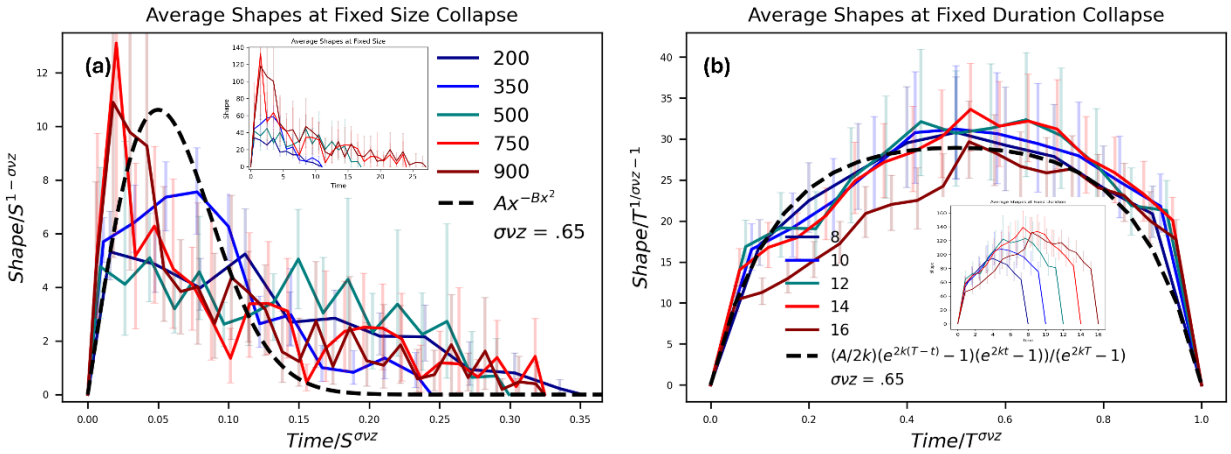


Figure 3.3: Scaling collapses of avalanches shapes. Scaling collapses of avalanche shapes at fixed a) size and b) fixed duration. The inserts show the unscaled shapes of each. The x and y axis for the shapes collapse at fixed duration (T) are t/T and shape/$T^{(1/\boldsymbol{\sigma\nu z})-1}$, respectively. The shapes at fixed size (S) collapse x and y axis are scaled by $t/S^{\boldsymbol{\sigma\nu z}}$ and shape/$S^{1-\boldsymbol{\sigma\nu z}}$, respectively. The value of σνz = .5 for the fixed size collapse and .6 for the fixed duration collapse. The black dashed line in (a) shows the predicted mean field scaling collapse function ($Ax^{(-Bx^2)}$(A = 175 and B = 90)). The black dashed line in (b) shows the predicted mean field scaling collapse function $\frac{A}{2k}\frac{e^{2k(T-t)-1}(e^{2kt}-1)}{e^{2kT}-1}$ plotted for (A = 210, k = 3, and T = 1)). The virus mutation rate is $1.17E^{-6}$. The system size of N hosts = 316,228.

It's important to know when in the time series the avalanches whose sizes are within the scaling regime are occurring. **Figure 3.4** shows that the avalanches within the scaling regime are occurring during host dominant regimes which is when the host is suppressing viral outbreaks (flatline on the host abundance vs time). The data also shows the avalanches are occurring as the virus population dies out at the end of the simulation. The last endemic finishes, and the hosts remain near carrying capacity with small drops in population as the virus dies out.

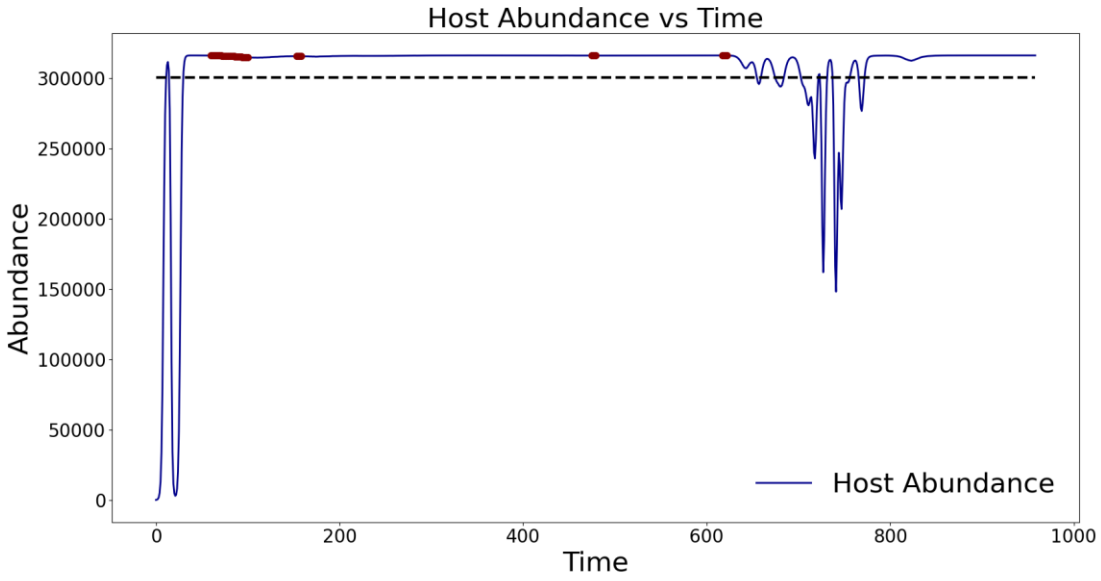


Figure 3.4: Scaling regime avalanches in time. The graph shows a time trace of the host abundance across time (blue line). The red dots indicate when avalanches occur, whose sizes fall within the scaling regime. The black, dashed line represents a host abundance at 95% of carrying capacity.

However, it is also worth noting how catastrophically large avalanches are distributed in the large avalanche regime comprised of system spanning events, that affect a macroscopic fraction hosts). **Figure 3.5** shows the statistical distributions for the duration versus size, size CCDF, and duration CCDF of the large avalanches. In this regime power law scaling is not necessarily expected because the system size sets a scale that can break the self similarity. Just to quantify the observations and highlight the difference to the actual scaling regime, we plotted approximate power laws next to the data. The roughly approximated power law exponent $\sigma\nu z$ from the duration versus size plot appears negative (-.5) showing that in this regime larger avalanches take less time. Like the small avalanche regime, the shapes of fixed sizes and durations approximately collapse (**figure 3.6**) on one another using the approximated $\sigma\nu z$ from the size vs duration plot of the large avalanches. The host abundance across time (**Figure 3.7**) shows that the avalanches in the large regime occur leading up to the largest outbreak within an endemic or

coming off the largest breakout within an endemic. Note that the exponents are not expected to match the mean field modal, but I plotted them to compare these large avalanches to the results in the small-avalanche scaling regime discussed earlier (see also **Table 3.2**).

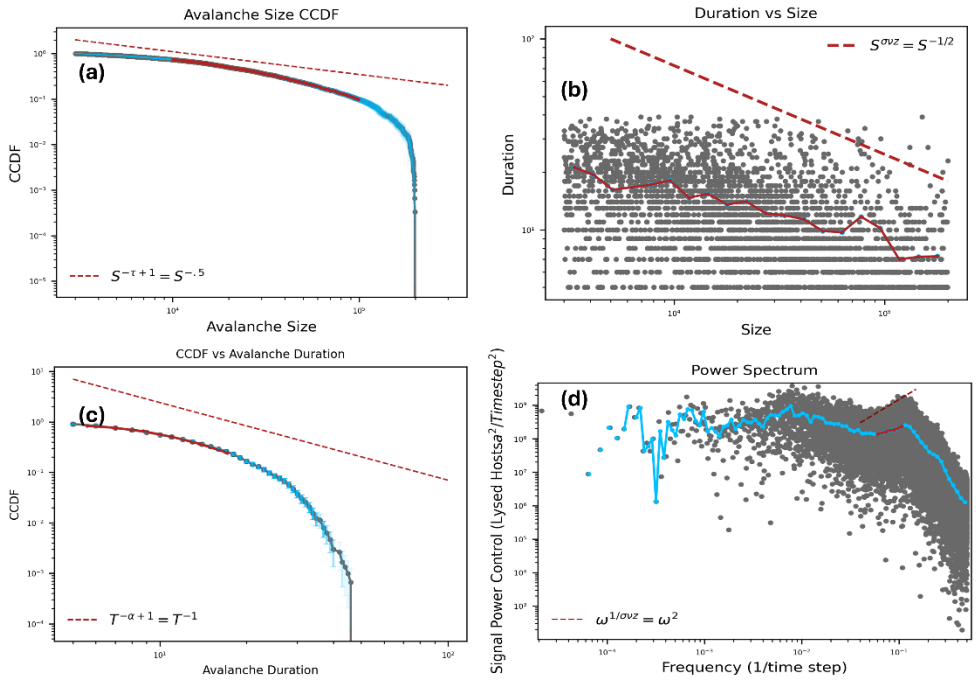


Figure 3.5: Large regime power laws: The statistics demonstrate a power law (red line) across the loglog plot of the a) CCDF of avalanche sizes, b) the avalanche size vs duration, c) CCDF of the avalanche duration and d) the power spectrum. Power law scaling is not expected in this regime because the finite system size sets a scale and breaks self-similarity. We just plot the approximate exponents σνz, τ, and α to highlight the differences compared to the true scaling regime that was discussed above for the small avalanches. Again, σνz, τ, and α give the power laws of the size versus duration, size CCDF, and duration CCDF, respectively. The system size was N hosts = 316,228 with viral mutation rate of $1.17E^{-6}$.

Exponents	$\boldsymbol{\tau}$ **(size dist.)**	$\boldsymbol{\alpha}$ **(duration dist.)**	$\boldsymbol{\sigma\nu z}$ **(size vs duration)**	$\boldsymbol{\frac{\tau - 1}{\alpha - 1} = \sigma\nu z}$
Host Avalanches	$\mathbf{1.31 \pm \frac{.23}{.25}}$	$\mathbf{1.08 \pm \frac{.54}{.08}}$	$\mathbf{-.28 \pm \frac{N/A}{.N/A}}$	$\mathbf{3.8 \pm \frac{10.0}{3.4}}$
Mean Field Theory (only listed for comparison -- MFT exponents should only apply to the scaling regime of the small avalanches discussed earlier, they are not expected to apply to this large avalanche regime.)	**1.5**	**2**	**.5**	**.5**

Table 3.2: Scaling exponent values for large regime. This table lists the values of each scaling exponent for the host avalanches. The mean field exponents are listed only for comparison. The mean field theory does not predict that these scaling exponents should apply in this regime, but rather that the large avalanches should have a different statistics than the small avalanches. These values were obtained using the slopes of the graphs in Figure 2 which use all spatial acquisition probabilities and a virus mutation rate of $1.17E^{-6}$. I used bootstrapping to calculate the exponents and error bars (Alstott et. al., 2014; Clauset et. at., 2009)

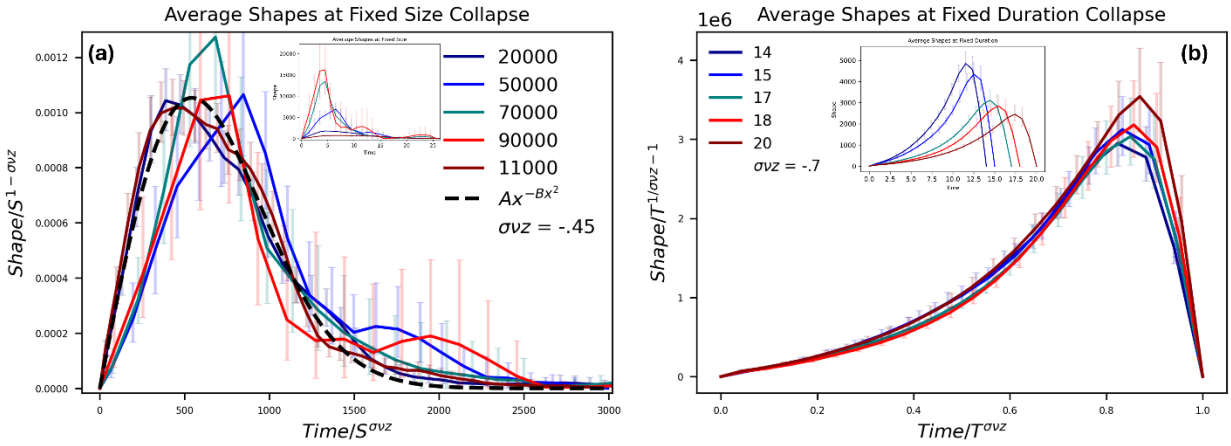


Figure 3.6: Approximate collapses of avalanches shapes in the large regime. Rescaling of the large avalanche shapes in the analogous way as was done in the scaling regime renders approximate collapses of the avalanche shapes at fixed a) size and b)

Figure 3.6 (cont.); fixed duration The insets show the unscaled shapes of each. The *x* and *y* axis for the shapes collapse at fixed duration (T) are t/T and $<V(t|T)>/T^{(1/\sigma\nu z)-1}$, respectively. The shapes at fixed size (S) collapse *x* and *y* axis are scaled by $t/S^{\sigma\nu z}$ and $<V(t|S)>/S^{(1/\sigma\nu z)-1}$respectively with $\sigma\nu z$ = -.45 for the fixed size collapse and -.7 for the fixed duration collapse. The black dashed line in (a) shows the predicted mean field scaling collapse function ($Ax^{(-Bx^2)}$(A = 3.2 * 10^{-6} and B = 1.7* 10^{-6})). The virus mutation rate $1.17E^{-6}$.

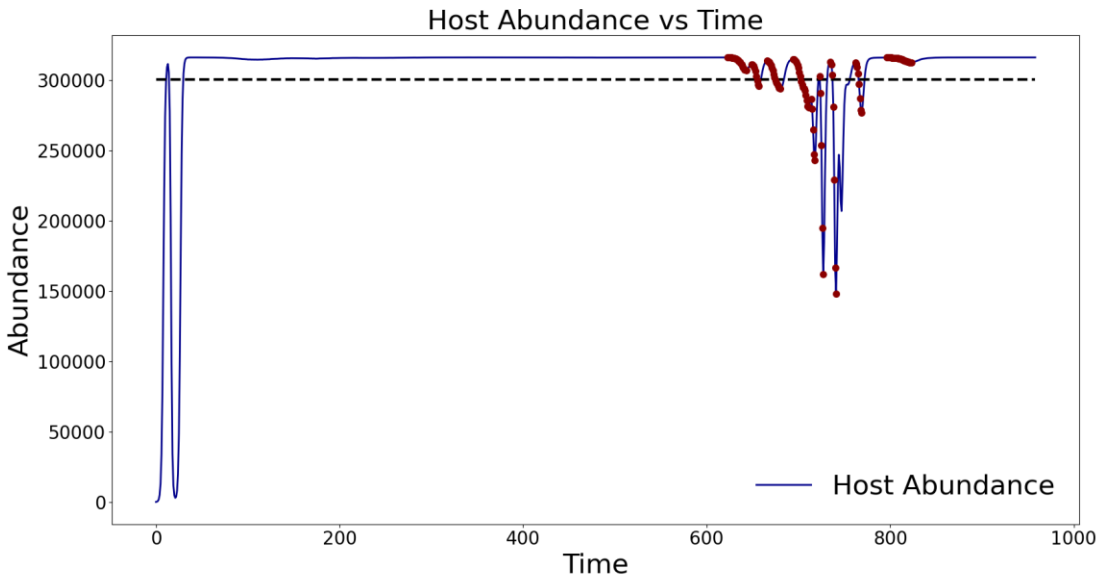


Figure 3.7: Large regime avalanches in time. The graph shows a time trace of the host abundance across time (blue line). The red dots are time points during avalanches within the large scaling regime. The black, dashed line represents a host abundance at 95% of carrying capacity.

Discussion

Interestingly, the two regimes play a role in identifying two key aspects of an endemic. The small regime has self-similarity and universality amongst avalanches within the scaling regime. These same avalanches are only present in times of virus suppression shown in figure 3.4. Interestingly, the large regime only shows approximate self-similarity and (but not necessarily full universality, because the finite size effects are expected to introduce a scale into the system) for

avalanches within a certain approximate scaling regime. In contrast to the small regime, the large regime avalanches are only present in times of macroscopically large host populations drops which are correlated to times of large virus population spikes. The large avalanches are part of an active and significant endemic.

A difference between the avalanche propagation in both regimes is shown in the shapes of the avalanche durations (Figure 3.2b and 3.6b). The small regime has avalanche duration shapes that are symmetrical whereas the large avalanches are leaning to the right, as has been seen in other systems where finite size effects are noticeable. If finite size effects would not be the leading cause then the right lean could also be indicative of inertia causing a slow increase of the avalanche velocity. The velocity then rapidly increases toward the end of the avalanche due to the large number of hosts lysed as time goes on. The large number of lysed hosts creates something of a feedback loop resulting in more hosts being infected. None the less, the shapes at fixed duration are different for the small avalanches in the scaling regime and the large avalanches. For data of real endemics the difference can help identify if the avalanche belongs to the large or small regime based on initial avalanche measurements.

Conclusion

The statistics of host population burst dynamics coexisting with bacteria with the capability of CRISPR acquired immunity are similar to previous MFT avalanche models. The simulations allowed for the analysis of criticality, self-similarity, and universality amongst viral burst avalanches spanning across a wide range of sizes. From the analysis we found the scaling exponents in the critical regime matched the exponents of MTF avalanche statistics. Understanding the statistics in the critical regime allows for the development of a simple model of the avalanches of host drops.

Chapter 4: The eerie agreement of flu and COVID-19 outbreaks with the model

Introduction

In chapter 2, I showed how a coarse-grained simulation representing host population drops due to an infectious virus follows the mean field model for magnet moments. Now, let's think about people getting infected by a virus. In this scenario, a person would be considered a host similar to a bacteria host. However, a human may not die but get sick. A healthy human will have a spin of "down" whereas the same person will have a spin of "down" when sick. With that in mind, infectious disease case statistics should also follow the mean field model for magnetic moments. Luckily there are statistics readily available through the W.H.O. and C.D.C.

Avalanches can be within the scaling regime and outside of the scaling regime as shown in chapter 2. These different areas produce avalanches of different shapes. Theoretically, we would be able to determine if an avalanche size falls within or larger than the scaling regime. We could determine where the avalanches fall based on the avalanche's shape. For example, a system spanning avalanche will have a more Gaussian looking shape. In contrast, avalanches within the scaling regime are predicted by the model to fluctuate heavily and their average shape (for a given avalanche size) follows the form $t \exp(-A t^2)$, where t is the time and A is a constant. If infectious diseases follow this same model, then we could use the avalanche shapes to determine the severity of a viral outbreak. The outbreaks spanning the human population (pandemic) will look more Gaussian whereas smaller outbreaks will fluctuate heavily in time and only on average follow the $t \exp(-A t^2)$ shape. Distinguishing between the two types of avalanches is useful to determine safety measures to keep people safe and to allocate our resources to fight an outbreak. In this chapter, I will investigate the dynamics of viral infections and their eerie similarities to the mean field model for magnetic moments.

Results

Despite its simplicity, the model is consistent with weekly global flu cases (data from the website of the World Health Organization). **Figure 4.1** shows the power laws associated with the size CCDF, size vs duration, duration CCDF, and the power spectrum. The dark red line is associated with the theoretical scaling exponent for each and is consistent with the slope of the power law for each figure. The scaling exponents σνz, τ, and α match closely with the theoretical mean field exponents (**Table 4.1**). Finally, the shapes at fixed size (**Fig. 4.2a**) and duration (**Fig. 4.2b**) graphs collapse one unto another when using the appropriate scaling equations and the measured value for σνz. The scaling collapse function is plotted on top of the shape collapse at fixed size for the flu data. The model prediction is consistent with the shapes of the flu avalanches. For clarity the figure does not show statistical error bars here, but instead, in **Figure 4.2** we show the shapes with error bars.

COVID-19 cases at the University of Illinois at Urbana-Champaign (UIUC) are also consistent with the model. UIUC reported daily COVID-19 tests that were used as a velocity trace to calculate the COVID-19 avalanches across the campus. The global COVID-19 data from the W.H.O. is shown in **Figure 4.5**. We focus on the University data here over the global data, because there is too much variation in the social, phycological, and physical variables between countries that would need to be accounted for with global data. However, UIUC results had consistent parameters of total population tested and testing rate. The results show a scaling regime in the size CCDF, size vs duration, duration CCDF, power spectrum, and max velocity vs size (Figure 4.3) that are consistent with the mean field exponents. There were not enough data points to get a good average from the shapes at fixed size and duration. The exponents are represented in **Table 4.1.** The model also works with Global Tuberculosis data (**Figure 4.7**) however, the W.H.O collected

data yearly so not enough data was available to extract the tuberculosis outbreak durations. Therefore, here only the outbreak sizes were analyzed.

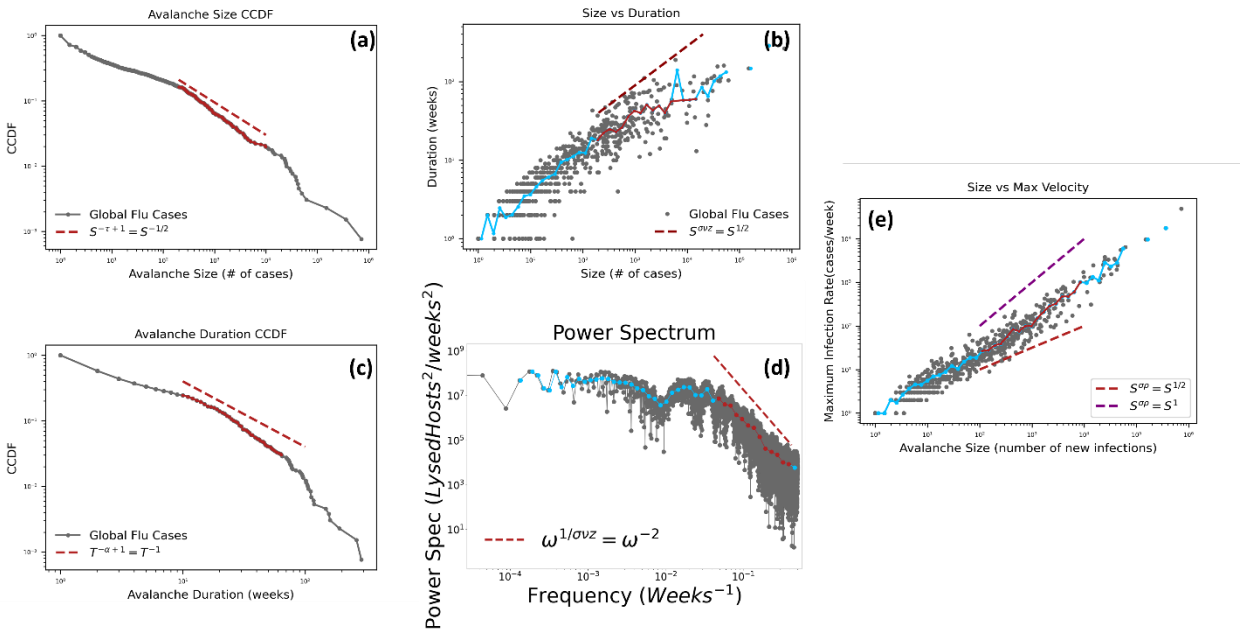


Figure 4.1: Global weekly flu cases: Scaling regimes from the (a) Outbreak size CCDF, (b) Size and duration with the binned size vs duration on overlayed (number of size bins = 50), (c) duration CCDF, (d) power spectrum and (e) outbreak size vs max velocity (*i.e.* maximum number of new infections during the outbreak). The red line indicates the powerlaw predicted by the mean field model. Exponents listed in the table are derived from the power laws of the size vs duration, size CCDF, and duration CCDF, respectively. Data used was from the World Health Organization **[3].**

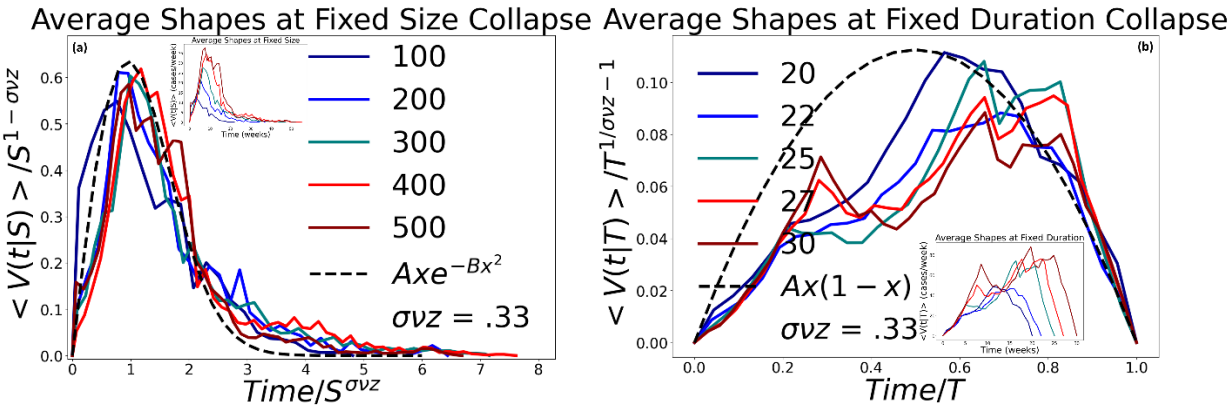


Figure 4.2: Average temporal profiles of flu outbreaks (avalanches shapes) for (a) fixed outbreak sizes and (b) fixed outbreak durations for weekly flu cases from each country. The insets show the unscaled profiles at different sizes (a) and different durations

Figure 4.2 (cont.); (b). The main panels show the scaling collapses. The black dashed line in (a) shows the predicted mean field scaling collapse function (Axe^{-Bx^2}plotted (A = 1 and B = 2). The black dashed line in (b) shows the predicted mean field scaling collapse function (Ax(1−x) plotted (A = 2.2)). The value σvz = .4 was used for both scaling equations to get the collapse. Data used was from the **Figure 4.2 (cont.);** World Health Organization **[3]**. The asymmetry to the right (b) indicates that the dynamics of flu may have a slight "inertia-like effect", or finite size effects could also tilt the shapes this way **[5]**.

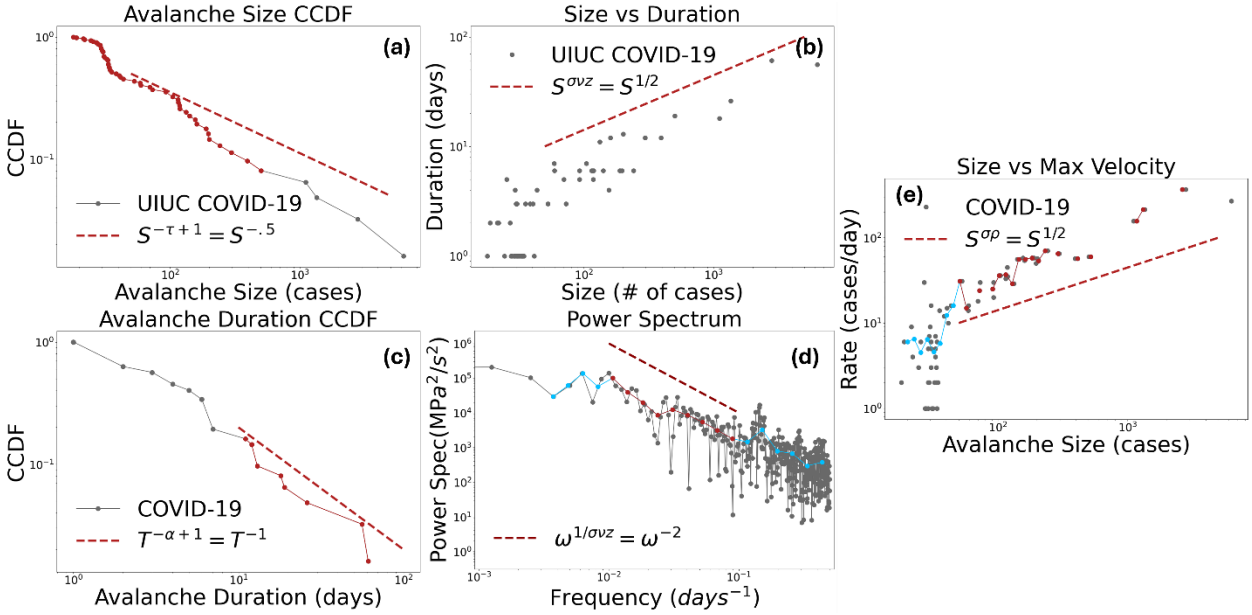


Figure 4.3: Daily COVID-19 cases from the University of Illinois at Urbana Champaign: Scaling regimes for the (a) Outbreak size CCDF, (b) outbreak duration versus outbreak size with the binned size vs duration on overlayed (number of size bins = 50), (c) outbreak duration CCDF (d) power spectrum of the time trace of daily case numbers and (e) outbreak size vs maximum "velocity" (*i.e.* maximum number of daily cases) for daily COVID-19 cases. A red line indicates each power law. Exponents in Table 1 are derived from the power laws of the size vs duration, size CCDF, and duration CCDF, respectively. Data used was from the University of Illinois **[6]**.

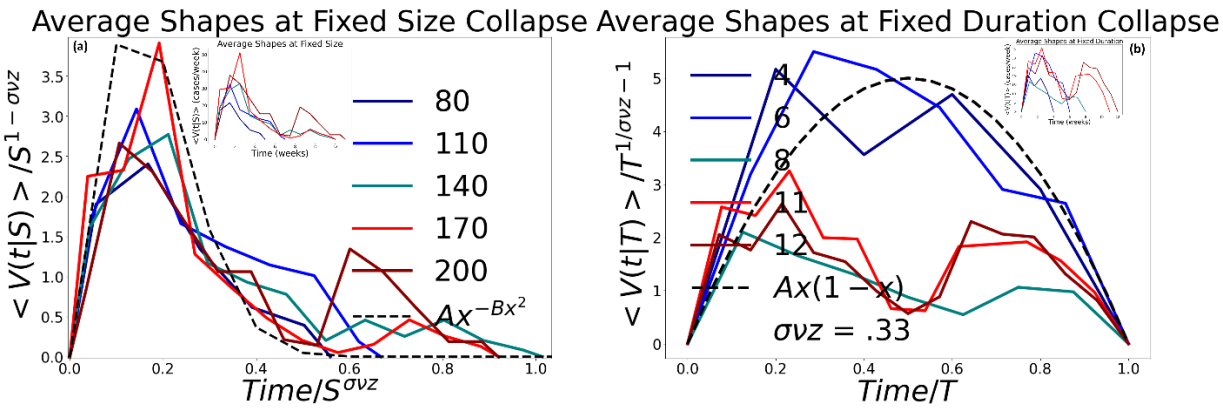


Figure 4.4: Average temporal profiles of flu outbreaks (avalanches shapes) for (a) fixed outbreak sizes and (b) fixed outbreak durations for weekly flu cases from each country. The insets show the unscaled profiles at different sizes (a) and different durations (b). The main panels show the scaling collapses. The value σvz = .5 and .4 for the fixed size and fixed duration scaling equations to get the collapse. The flattening of the shapes in (b) for the longer avalanches suggests that these outbreaks may have been resulting from temporal overlaps of several independent outbreaks. Data was taken from S.H.I.E.L.D.

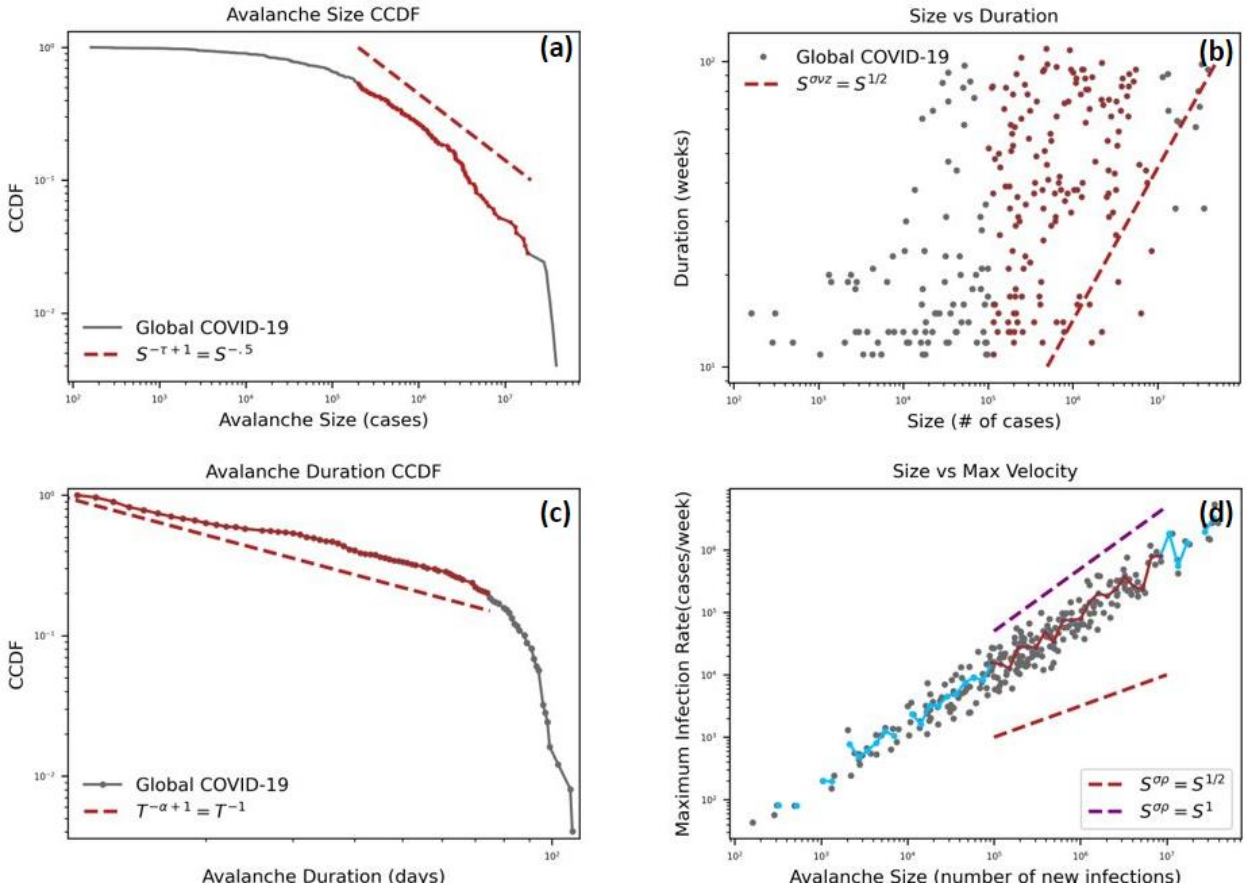


Figure 4.5: World-wide COVID-19 data: Scaling regimes from the (a) Size CCDF, (b) Size and duration, (c) duration CCDF, (d) outbreak size vs max velocity (*i.e.* maximum new infection rate). The red line indicates the integrated and non-intergated

Figure 4.5 (cont.); powerlaws that are predicted by the simple mean field theory. The exponents in table SI1 are derived from the powerlaws of the size vs duration, size CCDF, and duration CCDF, respectively. Data is weekly COVID-19 cases categorized by the W.H.O https://data.who.int/dashboards/covid19/data?n=c

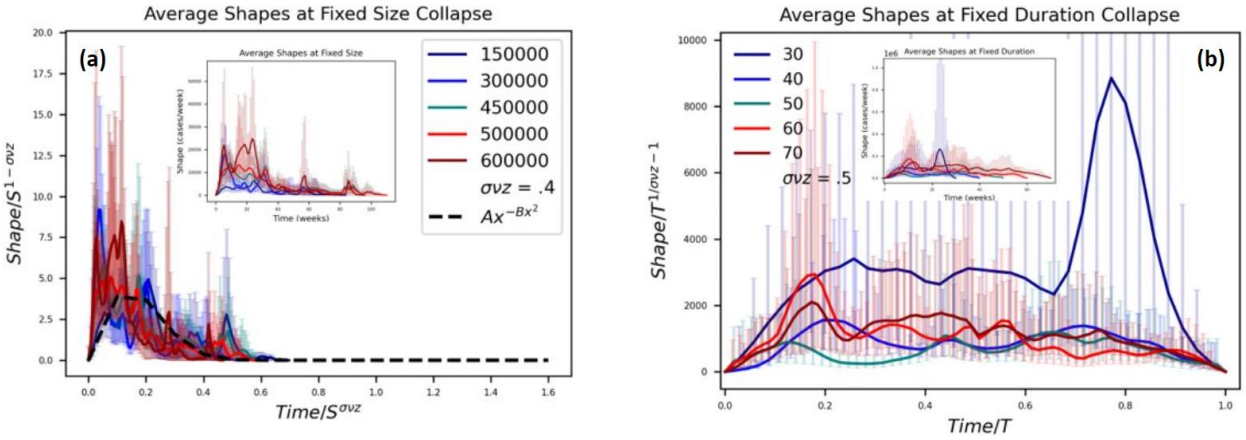


Figure 4.6: COVID Data worldwide: Avalanches shapes for (a) fixed outbreak sizes and (b) fixed outbreak durations for weekly flu cases from each country. Various fixed sizes of avalanches were used for each curve in the insert. The main figure shows the collapse of these curves when the x and y-axis are scaled via the appropriate scaling exponents. Scaling collapse function plotted (black, dashed line for nonuniversal values of the constants A = 50 and B = 25). The value σvz = .4, .5 was used for the size and duration scaling equations, respectively, to achieve the collapses. Data used was from the W.H.O. website https://data.who.int/dashboards/covid19/data?n=c

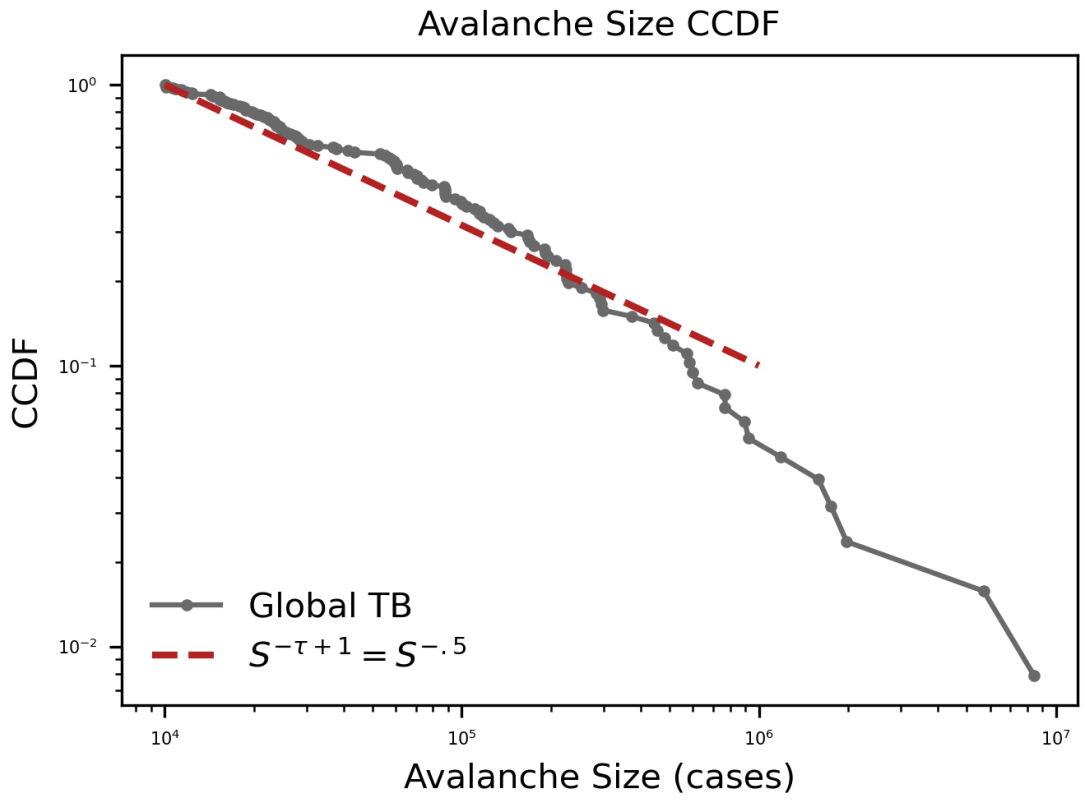


Figure 4.7: Worldwide Tuberculosis data: Scaling regimes from the Size CCDF. The red line indicates the mean-field power law. The exponent τ is derived from the power law of the size CCDF. The data is derived from yearly Tuberculosis cases categorized by the W.H.O. (https://www.who.int/teams/global-tuberculosis-programme/data#csv_files)

Table 4.1	$\boldsymbol{\tau}$	$\boldsymbol{\alpha}$	$\boldsymbol{\sigma\nu z}$	$\boldsymbol{\frac{\tau - 1}{\alpha - 1} = \sigma\nu z}$
Flu	$1.41 \pm \frac{.12}{.09}$	$1.71 \pm \frac{.54}{.45}$	$.33 \pm \frac{.04}{.06}$	$.57 \pm \frac{.97}{.26}$
UIUC COVID-19	$1.66 \pm \frac{.51}{.54}$	$1.74 \pm \frac{.50}{.28}$	$.60 \pm \frac{.25}{.16}$	$.87 \pm \frac{.66}{.86}$
Global COVID-19	$1.68 \pm \frac{3.29}{.57}$	$1.22 \pm \frac{.37}{.22}$	$.563 \pm \frac{.058}{.064}$	$.362 \pm \frac{2.34}{.352}$
Global Tuberculosis	$1.43 \pm \frac{.33}{.21}$	N/A	N/A	N/A
Mean Field Theory	.5	2	1.5	.5

Table 4.1: Theoretical and measured scaling exponents: Listed theoretical and measured values for the scaling exponents τ, α and $\sigma\nu z$ global COCID-19 outbreaks. The COVID-19 scaling exponents derived from their outbreaks match the mean field theoretical exponents. This shows that the infectious disease data match the model.

Discussion

The simulation results mimic real world infectious diseases as shown by the weekly global flu outbreaks and daily COVID-19 infections at the University of Illinois at Urbana-Champaign (UIUC). The number of flu cases is self-reported weekly by individual countries. The UIUC campus was tested twice weekly, and the positive COVID-19 cases were reported daily. The scaling exponents extracted from the outbreak size CCDF, size vs duration, duration CCDF, power spectrum, and maximum velocity (daily new infections) vs outbreak size are close to those of the simple mean field exponents (**Table 4.1**)**.** These exponents were obtained from the scaling collapses of the size CCDF, shapes at fixed size, and shapes at fixed duration collapse **(Figure 4.2**) for the flu data. From these results, we conclude that the outbreak statistics and dynamics of the flu and COVID-19 outbreaks can to good approximation be modeled by the simple mean field theory avalanche theory. The agreement is surprising because the mean field theory assumes a well-mixed situation, while, during the COVID pandemic, great care was taken to ensure social distancing.

The agreement of the model predictions with flu and COVID-19 infections shows how even a simple model can mimic the main trends of the statistics of complicated infectious diseases. There are many different infectious diseases with different infection rates and degrees of severity that will be studied in future work. The coarse-grained model is useful for gaining intuition for how to best organize the observational data. It is likely that variants of this model work also for bloodborne and contact specific infectious diseases in addition to the airborne diseases studied

here. It might be that other transfer mechanisms may change the effective dimensionality of the model from mean-field-like for airborne diseases (well mixed), to a two-dimensional case for example. The model would predict that for a disease with essentially two-dimensional dynamics, large outbreaks (avalanches) are generally more strongly suppressed than in the well mixed case. Also, the scaling forms and scaling exponents in two dimensions would be different (**Dahmen et. Al., 1996; Perkovic et. al., 1999; Kuntz et. al., 2001),**

The model can be applied to other infections as well. **Figure 4.6** shows power law scaling in plots of the outbreak size CCDF, outbreak size vs duration, duration CCDF, power spectrum of the time trace of new infections, and max new infection rate vs avalanche size for weekly new COVID-19 infections from each country. The avalanches were calculated from each country's reported COVID-19 cases. Then, all the avalanches were organized into one data set to extract the statistics below. The scaling exponents are similar to the theoretical mean field values shown in **Table 4.1**. The simulated results mean field scaling exponents for the case where the size distributions have been integrated over all virus mutation rates. Infectious diseases are made up of viral infections where the viruses specific to the disease have a unique mutation rate. For this reason, the scaling exponents for the infectious disease data may not be the exact same as the scaling exponents for the simulations at the slowest possible adiabatic rate. Furthermore, for simulations in finite systems the mean field exponents and mean field integrated exponents deviate slightly based on which method is used. The shapes at fixed size, fixed size collapse **(fig 4.5a)**, fixed duration, and fixed duration collapse **(fig. 4.5b)** use values for the scaling exponents that are consistent with the mean field value of $\sigma\nu z$. Each country dealt differently with many factors politically, socially, mentally, and physically regarding COVID-19. For that reason, the varying population, population density, mask mandates, work from home mandates etc. could have

differently affected the statistics for different countries. More research needs to be done in the future to appropriately account for the difference in these conditions in each country.

Conclusion

The scaling exponents for flu and COVID-19 are well approximated by the mean field model for a well-mixed system. Large outbreaks can be detected for sufficiently large initial infection rates, as is easily seen from the temporal profile shape <V(t)> at various sizes larger than the events in the scaling regime. Their shapes resemble a Gaussian profile compared to the exponential profile of avalanches within the scaling regime (**Figure 2.9).** The model predicts many statistical measures that could be compared to other diseases. There are less coarse-grained models for viruses and host population dynamics using Lotdka-Volterra equations (see for example (Pilsof et., al.)). Pilsof et., al. model the dynamics using many different virus mutation rates and spacer acquisition probabilities (chapter 3), so that their model has many more parameters than the simple coarse-grained model studied here. The changes in virus mutation rate and spacer acquisition probabilities require significant computational power to identify the critical point in their model because many parameter sweeps would be necessary. The simple mean field model studied here only uses one effective spacer acquisition probability as J_{inf} and integrates over all virus mutation rates (H(t)) (chapter 2). This coarse-grained model could be used to gain intuition for the more detailed simulations of complex models. The model also can be used to develop strategies to reduce the number of infections and develop a systematic approach for computational power allocation.

Appendix A: Recovery of HLMVEC junction strength after glycolysis is restored.

The data from chapter one of the thesis was compelling, but there was a lingering question on whether the HLMVECs died during the PFK-15 treatment. To test this concern, cells were treated with PFK-15 for 45 minutes. Then the media containing PFK-15 was removed, and the cells were washed with PBS. New media was added to the cells and the cells were incubated overnight to allow the production of ATP to recover.

Data was collected from TFIMB measurements for the average junction tension (**Figure A1**) and the junction tension increased from the recovered cells compared to the inhibited cells. The junction forces of the "washed off" cells show a recovery of a junction force from each cell applied to one another. The traction stress also increased as shown by the red shift of the heatmap of the washed off cells compared to the inhibited cells and the histogram of traction stresses **(Figure A2).**

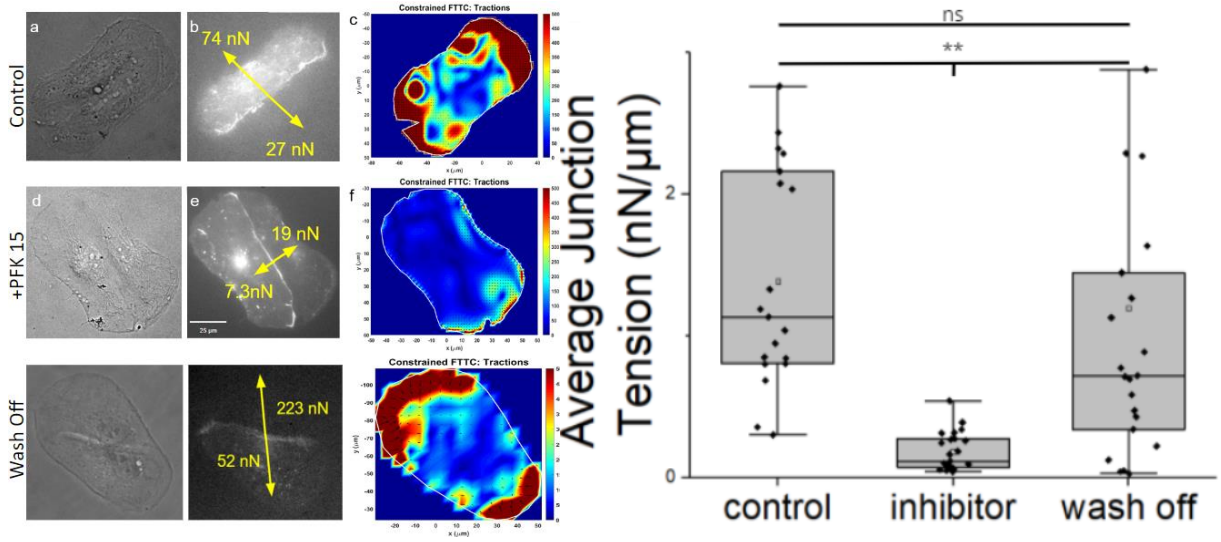


Figure A1: Traction Force Imbalance (TFIMB) measurements of HLMVEC pairs. Substrates were 40 kPa gels. DIC Images of untreated (control), PFK15 treated (+PFK15) cell PFK 15 washed of and with the control media (wash off) pairs on fibronectin patterns. Fluorescence image of control (control) and PFK15 treated (+PFK15) cells expressing GFP-VE-Cadherin and the determined forces exerted on the junction by each cell. Heat maps of traction stresses exerted by control, PFK15 treated, and washed

Figure A1 (cont.); off HLMVEC doublets seeded on stamped fibronectin patterns of tangential circles. (h) Average junction tension of control, +PFK15, and washed off cells.

The recovered cells peak is around 150 Pa and has a large tail the extends out to 2,500 Pa. These results show that the traction forces produced from the cells increased once PFK-15 was replaced with the controlled media as the inhibited cells have a large peak of traction stress at 25 Pa. The TFIMB and traction stress results show that the cells were not dead at the time of inhibited measurements and the cell is able to re-gain their ability to produce contractile forces on their substrate. The wind rose plots of the recovered cells show more traction forces angled with relation to the cell-cell junction uniformly toward the center in contrast. to the non-uniformity of the +PFK 15 cells. The results of the wind rose plots demonstrate that the cells contractile forces point radially inward as once media is switched back to the control media after PFK 15 treatment.

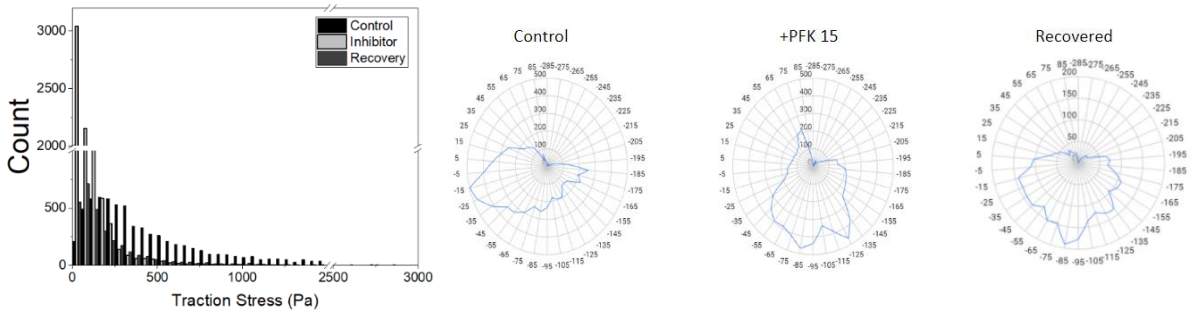


Figure A2: Traction stress distribution and orientation. Wind rose plots of the angles of traction force vectors relative to the junction normal in control, inhibited, and recovered HLMVEC doublets. Histogram of the magnitudes of measured traction stresses exerted by cells doublets on printed fibronectin patterns on 40 kPa gels. Black bars indicate control cells, dark grey are the recovered cells, and gray bars were obtained with inhibitor treated cells.

The recovery of contractile forces leads to the theory that the subcellular protein, actin, has re-polymerized. **Figure A3** shows HLMVECs stained with phalloidin and undergoing fluorescent imaging with a wide view microscope. The inhibited cells do not show f-actin fibers but diffuse actin staining across the cells (as shown in chapter 1). Yet, the cells that had been given the control

media post PFK 15 treatment show to have f actin fibers spanning the cells. The conclusion from these results is actin fibers can re-polymerize post glycolysis inhibition.

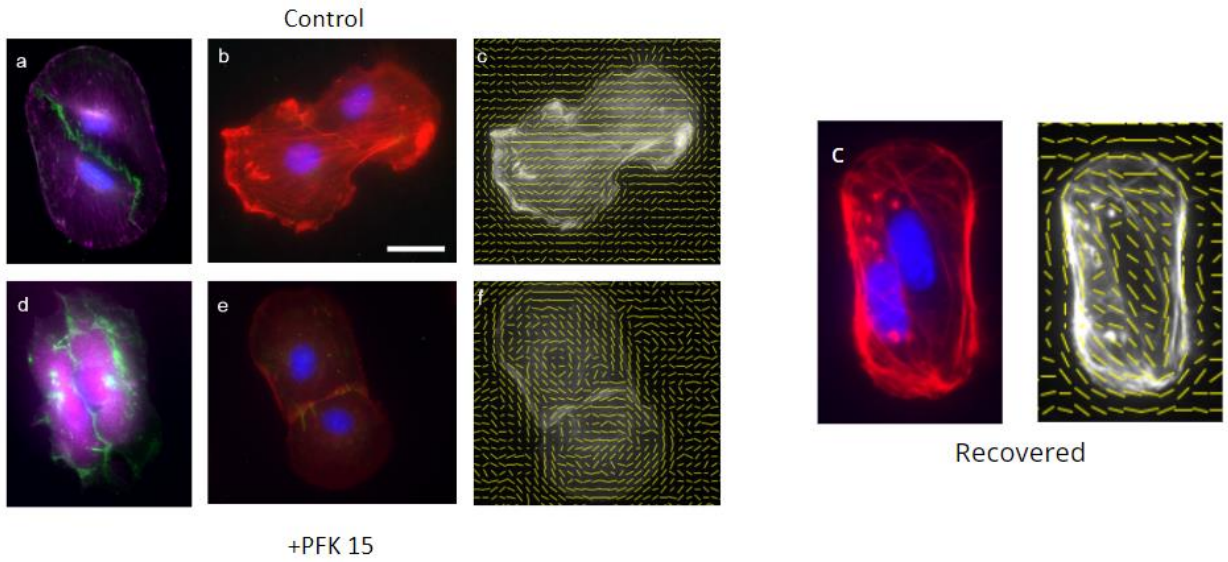


Figure A3: Inhibiting glycolysis disrupts focal adhesions and F-actin: Merged immunofluorescence images of paxillin (purple) and VE-cadherin (green) in untreated and PFK15 treated HLMVEC doubles. Merged immunofluorescence images of actin (red) and VE-cadherin (green) in untreated, PFK15 treated, and recovered HLMVEC doublets. Nuclei are stained with DAPI. G/f/H) Vector maps showing the directionality of the actin fibers in untreated, PFK15 treated, and recovered HLMVECs. In all images, cells were on micro contact printed fibronectin on 40 kPa gels. Scale bar = 25 µm

APPENDIX B: Materials and Methods from Chapter 1

Cell lines, adhesive protein and inhibitor:

Human Lung Microvascular Endothelial Cells (HLMVECs) obtained from Cell Applications Inc. were maintained in growth medium EGM-2MV (Lonza) medium supplemented with 10% (v/v) fetal bovine serum (FBS, Sigma – F0926-500ML) and 1% (v/v) penicillin– streptomycin (Corning Cell Grow, Manassas, VA). Glycolysis was inhibited, by incubating cells with 10 μM PFK15 (Cayman Chemical - #17689) for 40 min, prior to measurements. Cells adhered to substrates coated with fibronectin (Sigma SIGMA-ALDRICH INC. - FC010-5MG) diluted 10x (0.1mg/mL) with phosphate buffered saline (PBS) (Corming, #21-040-CV). HEK293T and HEK293AD cells were cultured in DMEM (Corning) supplemented with 10% FBS and 1% PenStrep (Gibco).

Preparation of polyacrylamide gels:

The 35 mm diameter glass bottom dishes with13 mm wells (Cell E&G - #GBD00001-200) were treated with 0.1M NaOH in order to make the glass hydrophilic. Then the glass was treated with APTMS (Sigma), washed twice, and treated with (v/v) 0.5% gluteraldyhyde (Sigma) to covalently link the polyacrylamide gels to the glass. The Young's modulus of the polyacrylamide gels used for the experiments was 40 kPa. Gels were prepared with deionized (DI) water, 8% acrylamide, 0.48% bis acrylamide, and 2μL of 0.2μm red fluorescent beads (Life Technologies Corporation - F8810) (Engler et al., 2004). The latter beads are used as fiducial markers for traction force measurements. Crosslinking agent ammonium persulfate (10% w/v APS, Sigma Aldrich) and catalyst N,N,N′,N′-tetramethylethylenediamine (0.5 μL - TEMED, Sigma Aldrich) was used to initiate the reaction. Next, 20 μL of the gel mixture was pipetted onto the center of each glass bottom dish and covered with a 12 mm diameter glass cover slip (EMS Acquisition Corp / Electron Microscopy Sciences - 72231-01). After curing gels for 30 minutes at 37°C, the gels were

submerged in DI water to keep the gels hydrated and to remove any unreacted chemicals. Gels were stored at 40^{o}C for up to 2 weeks.

Preparation of PDMS stamps:

Polydimethylsiloxane (PDMS) stamps were created to microcontact print fibronectin in patterns of tangential 1600 μm^2 circles on 40 kPa gels. To create the PDMS stamps, 10:1 Sylgard 184 Silicone Elastomer Base, part B, (Sylgard^{tn} 4019862) Sylgard 184 Silicone Elastomer Curing Agent, part A, respectively. The mixture was poured over a negative master and put into a decanter for 1 hour, to remove bubbles. Then the mixture was cured for 15 minutes in a 108^{o} C oven. The PDMS was peeled from the master and cut into individual stamps (Qin et al., 2010)

Viral Infections

Plasmid Constructs: The FRET-based VE-Cadherin tension sensor (VE-CadTS: (Addgene Plasmid# 45848) was used to measure tension on VE-cadherin bonds. For expression in endothelial cells, both constructs were cloned into the PLJM1-EGFP lentiviral vector (Addgene, Plasmid #19319). The inserts were amplified by PCR, using the Q5 Polymerase (New England Biolabs) and ligated into the digested PLJM1 vector at AgeI and EcoRI restriction sites. Cloning resulted in the removal of EGFP from the vector backbone. Successful cloning was confirmed by Sanger sequencing. Additional confirmation was obtained by observing the correct reporter localization at cell-cell junctions, as visualized under a confocal fluorescent microscope.

Virus Production: Lentivirus was produced to express VE-CadTS and VE-CadTL in HLMVECs. To generate the virus, VSV-G (lentiviral envelope expressing vector, Addgene, Plasmid# 12259), psPAX2 (viral packaging vector, Addgene, Plasmid# 12260) and either PLJM1-VE-CadTS or PLJM1-VE-CadTL were co-transfected into HEK293T cells using JetPrime (Polyplus) according

to the manufacturer's protocol. The virus-containing supernatant was harvested 48 hours after transfection, and the viral particles were precipitated using Lenti-X concentrator (Takara Bio) following the manufacturer's protocol. Low passage HLMVECs were transduced with concentrated virus, along with 10µg/mL Polybrene (Millipore Sigma). Expression was observed 2-7 days following infection.

Adenovirus was used to express VE-Cadherin-GFP in HLMVECs (Shaw et al, 2001). The virus was amplified in HEK293AD cells, by transducing the cells and collecting the virus containing supernatant, once the cytopathic effect was complete. The viral particles were then purified using the Adeno-X Maxi Purification Kit (Takara Bio) as per the manufacturer's instructions. Low passage HLMVECs were transduced with the virus, and expression was observed 2-3 days following infection.

Viral infections for GFP-VE-cadherin and VE-CadTS expression in HLMVECs:

Viral infections were performed in accordance with established BSL2 protocols. HLMVECs were cultured in Human Microvascular Endothelial Cell Medium (HLMVEC Cell Applications) with 10% FBS and 1% penicillin-streptomycin. Cells grew to 60% confluence in 6-well plates before treating with the viral infections. The VE-CadTS and polybrene were added to the cells for the lentiviral infection whereas just the VE-cadherin GFP virus was added for the adenoviral infection. Cells were incubated with virus at 37°C under 5% CO_2 for 12-18 hours, before the medium was changed to virus-free medium. Cells were then cultured for 2 additional days, to allow for fluorescent protein expression.

METHODS:

Micro contact printing fibronectin patterns on polyacrylamide gels:

PDMS stamps were submerged in ethanol and sonicated for 5 minutes and plasma treated (Harrick Plasma - PDC-32G) to make the stamps hydrophilic. Fibronectin (0.1mg/mL) was pipetted onto the stamps and incubated for 1 hour, to allow fibronectin to adsorb on the stamps. Fibronectin was immobilized on Sulfo-SANPAH-activated (25mg/µl), 40 kPa gels by aspirating the PBS from the stamp and then stamping the adsorbed protein onto the activated gels. Gels were incubated at 37^{o} C before the stamps were removed. The gels with printed fibronectin patterns were then submerged in PBS and stored overnight at 4^{o} C.

Traction force microscopy:

Traction force microscopy (TFM) measurements used polyacrylamide hydrogels with Young's moduli of 40 kPa. Fibronectin (0.1mg/mL) was immobilized on Sulfo-SANPAH-activated gels, by incubation with 0.1 mg/ml fibronectin for 4 hr at 37°C (Sigma SIGMA-ALDRICH INC. - FC010-5MG). After incubation, gels were washed with 1x phosphate buffered saline (PBS). HLMVECs were seeded on the coated gels and incubated overnight at 37°C under 5% CO_2 to allow cells to reach mechanical equilibrium. HLMVECs were treated with PFK15 for 40 minutes. Control cells were untreated. The root mean squared (RMS) traction stress was determined, by comparing the position of the beads beneath the cells, before and after they were lysed. Traction force heat maps were calculated using a custom Matlab program (Butler et al. 2001), provided by Ning Wang (University of Illinois at Urbana-Champaign).

Traction force imbalance (TFIMB) measurements:

HLMVECs expressing the GFP-tagged VE-Cadherin were seeded at ~ 20,000 cells per gel and incubated at 37°C under 5% CO_2 overnight, to allow cells to reach mechanical equilibrium. Cells adhered to the stamped fibronectin patterns. We determined the traction forces exerted by the cells

(Butler et al. 2001). To determine the net force orthogonal to the cell-cell junctions, we used a custom Traction Force IMBalance (TFIMB) program, which was based on a prior report (Maruthamuthu et al., 2011). We calculated the unconstrained traction forces because the boundary conditions used to calculate constrained tractions force maps are incompatible with the TFIMB calculations. The junction tension (force/length) was calculated by dividing the average force orthogonal to the cell-cell contact by the junction length, which was visualized by imaging GFP-VE-cadherin.

Immunofluorescent imaging:

After cells were treated with PFK15 for 40 min, they were fixed with 2% paraformaldehyde for 15 min at 37°C under 5% CO_2. Control cells were not treated with inhibitor. Post fixation, cells were washed 3x with PBS for 5 minutes each on a slow rocker. Cells were permeabilized with (v/v) 0.1% 100x Triton mix (Sigma - T8787-100ML), washed 3x for 3 minutes with PBS on a slow rocker, and then incubated with a blocking buffer containing 1%(w/v) Bovine Serum Albumin (BSA), for 20 minutes. Then, the cells were treated with the primary antibodies and phalloidin for 1 hour. Focal adhesions were visualized by immunostaining paxillin with primary mouse anti-paxillin antibody (1:200 BD Transduction Laboratories – 3108546-612405) and secondary antibody goat anti-mouse 488 (1:200 Invitrogen - 41819A). VE-cadherin was immunostained with rabbit anti-VE-Cadherin (1:200 Cell signaling Technologies - D87F2) followed by secondary, goat anti-rabbit Cy5 (1:200 Invitrogen - 1843847). F-actin was stained with phalloidin 488 (1:200 Invitrogen - A 12379). After labeling with primary antibodies, the cells were washed five times at 5 min per wash, and then treated with the secondary antibodies. Finally, the cells were washed 5 times at 5 min per wash, with the last wash containing DAPI at a (v/v) 1,000x dilution. Cells were then treated with mounting medium (Invitrogen - 2147630). Images

were acquired with a Zeiss Axiovert 200M wide view microscope at 40x with an oil immersion objective, using Carl Zeiss™ Immersol™ Immersion Oil (Fisher Scientific - 12-624-66A). VE-cadherin was imaged by illumination at 647 nm laser, and the paxillin and F-actin fluorescent labels were excited at 488 nm.

FRET measurements of junction tension:

FRET images of junctions between HLMVECs expressing the VE-cad-TS were imaged at 40x with a Zeiss Axiovert 200M wideview microscope. Regions of interest were imaged at each junction with a CFP (filter set 47, #000000-1196-682, excitation: BP 436/20, emission: 480/40) filter and a FRET (CYP/YFP filter set 48, 000000-1196-684, excitation: BP 436/20, emission: 535/30) base pair filter. Intensities at the junction with the CFP filter, FRET filter, and background were quantified in imageJ. The background of each was subtracted for both filters in ImageJ, by measuring the background-subtracted intensity at the junctions. FRET/CFP ratios were calculated using ImageJ. As increased FRET indicates a decrease in FRET, the inverse ratio, CFP/FRET was graphed. Thus, an increase in CFP/FRET indicates an increase in tension.

For the measurements, ~ 20,000 cells were seeded onto each microcontact-printed, 40 kPa gel, prepared as described above. Cells were incubated overnight at 37°C under 5% CO_2, in order for cell doublets on the patterns to reach mechanical equilibrium. Before imaging, the cells were treated with the inhibitor, as described above. Control cells were not treated.

FRAP measurements of cadherin mobility:

We conducted fluorescence recovery after photobleaching (FRAP) measurements of VE-cadherin at junctions between HLMVECs in a confluent monolayer, using a Confocal – Zeiss Light

Scanning Microscope 880. GFP-VE-cadherin was expressed in HLMVECs following adenovirus infection, as described above. The GFP-VE-cadherin fluorescence in a region of interest of 17 μm^2 was photobleached for 500 ms at the maximum laser intensity of ~9.3 W/cm^2 and then allowed to recover with the laser at a reduced intensity of ~1.04 W/cm^2 for 10 minutes, with a recovery area of 29 μm^2. Intensities were measured continuously and recorded every 500ms during the measurements at 20x magnification and a wavelength of 488 nm. The intensities at the different time points were normalized relative to the pre-bleach intensity.

HLMVEC monolayers at ~85% confluence were cultured on fibronectin-coated, 40 kPa gels, as described above. Cells were incubated overnight at 37°C and under 5% CO_2. Prior to measurements, cells were treated with the inhibitor, as described above. Control cells were not treated.

Acknowledgments:

We thank Hyunjoon Kong and Paul Kenis (University of Illinois at Urbana-Champaign) for use of their plasma cleaners. We thank Austin Cyphersmith (Research Specialist at the Institute of Genomic Biology) for assistance with FRAP, FRET, and immunofluorescence imaging. This work was supported by AHA 18PRE34070092 to PG, and PO1 HL060678-16 to DEL and JR.

APPENDIX C: Complementary cumulative distribution plot of all spacer acquisition probabilities from chapter 3.

The results in Chapter 3 were from an aggregate of all parameter combos with the virus mutation rate 1.17 E^{-6}. **Figure C.1** shows the size CCDF of the individual spacer acquisition probabilities for the one virus mutation rate. The curves fall close to one another and have the same shape. For that reason, we decided to combine all the spacer acquisition probabilities into one dataset because it gives a lot more data points for each regime (small and large).

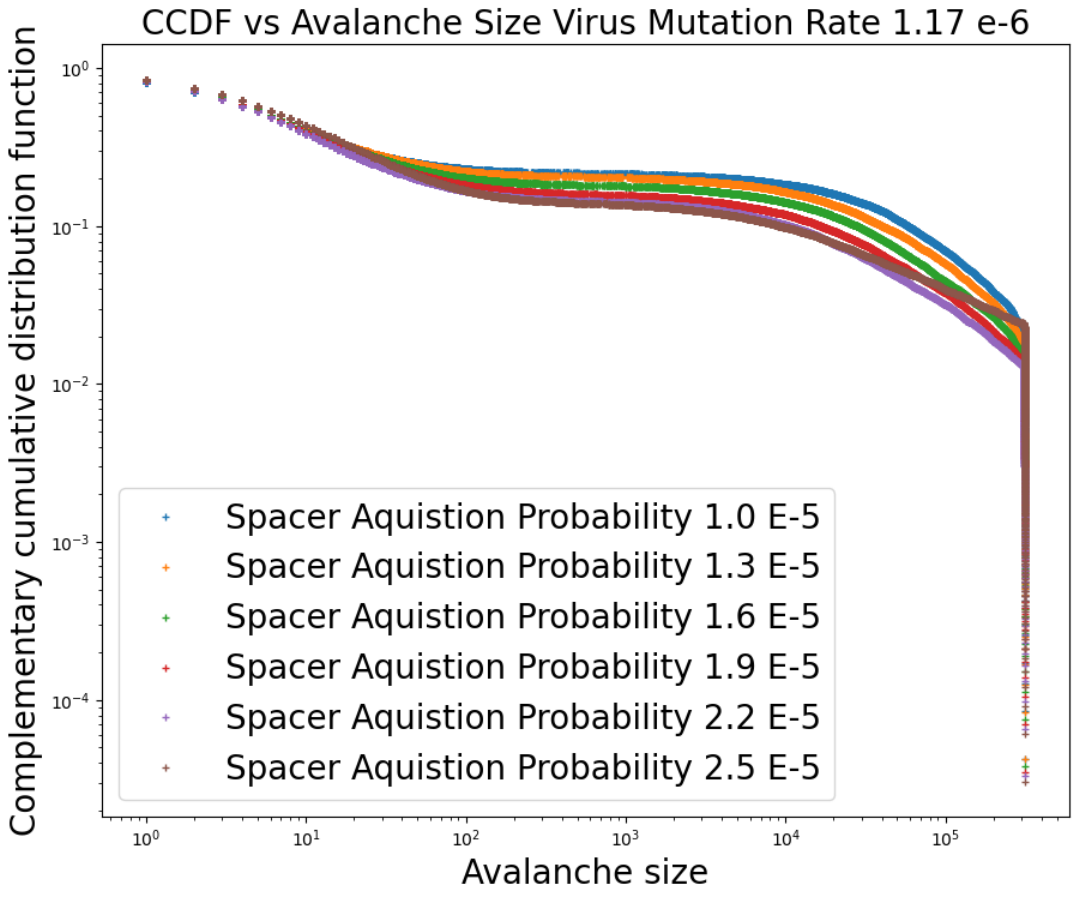


Figure C.1 Aggregate size CCDF of spacer acquisition probabilities: A CCDF of all spacer acquisition probabilities for virus mutation rate 1.17 E^{-6}.

References:

1. Barry, A. K., Tabdili, H., Muhamed, I., Wu, J., Shashikanth, N., Gomez, G. A., Yap, A. S., Gottardi, C. J., De Rooij, J., Wang, N., & Leckband, D. E. (2014). α-Catenin cytomechanics - role in cadherin-dependent adhesion and mechanotransduction. Journal of Cell Science, 127(8), 1779–1791. https://doi.org/10.1242/jcs.139014

2. Bays, J. L., Campbell, H. K., Heidema, C., Sebbagh, M., & Demali, K. A. (2017). Linking E-cadherin mechanotransduction to cell metabolism through force-mediated activation of AMPK. Nature Cell Biology, 19(6), 724–731. https://doi.org/10.1038/ncb3537

3. Bershadsky, A. D., Balaban, N. Q., & Geiger, B. (2003). Adhesion-Dependent Cell Mechanosensitivity. Annual Review of Cell and Developmental Biology, 19, 677–695. https://doi.org/10.1146/annurev.cellbio.19.111301.153011

4. Buckley, C. D., Tan, J., Anderson, K. L., Hanein, D., Volkmann, N., Weis, W. I., Nelson, W. J., & Dunn, A. R. (2014). The minimal cadherin-catenin complex binds to actin filaments under force. Science, 346(6209). https://doi.org/10.1126/science.1254211

5. Butler, J. P., Toli-Nørrelykke, I. M., Fabry, B., & Fredberg, J. J. (2002). Traction fields, moments, and strain energy that cells exert on their surroundings. American Journal of

Physiology - Cell Physiology, 282(3 51-3), 595–605. https://doi.org/10.1152/ajpcell.00270.2001

6. Cai, Y., Rossier, O., Gauthier, N. C., Biais, N., Fardin, M. A., Zhang, X., Miller, L. W., Ladoux, B., Cornish, V. W., & Sheetz, M. P. (2010). Cytoskeletal coherence requires myosin-IIA contractility. Journal of Cell Science, 123(3), 413–423. https://doi.org/10.1242/jcs.058297

7. Cao, J., & Schnittler, H. (2019). Putting VE-cadherin into JAIL for junction remodeling. Journal of Cell Science, 132(1). https://doi.org/10.1242/jcs.222893

8. Clem, B. F., O'Neal, J., Tapolsky, G., Clem, A. L., Imbert-Fernandez, Y., Kerr, D. A., Klarer, A. C., Redman, R., Miller, D. M., Trent, J. O., Telang, S., & Chesney, J. (2013). Targeting 6-phosphofructo-2-kinase (PFKFB3) as a therapeutic strategy against cancer. Molecular Cancer Therapeutics, 12(8), 1461–1470. https://doi.org/10.1158/1535-7163.MCT-13-0097

9. Conradi, L. C., Brajic, A., Cantelmo, A. R., Bouché, A., Kalucka, J., Pircher, A., Brüning, U., Teuwen, L. A., Vinckier, S., Ghesquière, B., Dewerchin, M., & Carmeliet, P. (2017). Tumor vessel disintegration by maximum tolerable PFKFB3 blockade. Angiogenesis, 20(4), 599–613. https://doi.org/10.1007/s10456-017-9573-6

10. Conway, D. E., Breckenridge, M. T., Hinde, E., Gratton, E., Chen, C. S., & Schwartz, M. A. (2013). Fluid shear stress on endothelial cells modulates mechanical tension across VE-cadherin and PECAM-1. Current Biology, 23(11), 1024–1030. https://doi.org/10.1016/j.cub.2013.04.049

11. Daneshjou, N., Sieracki, N., van Nieuw Amerongen, G. P., Schwartz, M. A., Komarova, Y. A., & Malik, A. B. (2015). Rac1 functions as a reversible tension modulator to stabilize VE-cadherin trans-interaction. Journal of Cell Biology, 208(1), 23–32. https://doi.org/10.1083/jcb.201409108

12. De Bock, K., Georgiadou, M., Schoors, S., Kuchnio, A., Wong, B. W., Cantelmo, A. R., Quaegebeur, A., Ghesquière, B., Cauwenberghs, S., Eelen, G., Phng, L. K., Betz, I., Tembuyser, B., Brepoels, K., Welti, J., Geudens, I., Segura, I., Cruys, B., Bifari, F., … Carmeliet, P. (2013). Role of PFKFB3-driven glycolysis in vessel sprouting. Cell, 154(3), 651–663. https://doi.org/10.1016/j.cell.2013.06.037

13. Dudek, S. M. Garcia, J. G. N. (2001). Cytoskeletal regulation of pulmonary vascular permeability: Journal Applied Physiology 91(4), 1487-1500 https://doi.org/10.1152/jappl.2001.91.4.1487

14. Fang, Y., Wu, D., & Birukov, K. G. (2019). Mechanosensing and mechanoregulation of endothelial cell functions. Comprehensive Physiology, 9(2), 873–904. https://doi.org/10.1002/cphy.c180020

15. Gawlak, G., Tian, Y., O'Donnell, J. J., Tian, X., Birukova, A. A., & Birukov, K. G. (2014). Paxillin mediates stretch-induced Rho signaling and endothelial permeability via assembly of paxillin-p42/44MAPK-GEF-H1 complex. FASEB Journal, 28(7), 3249–3260. https://doi.org/10.1096/fj.13-245142

16. Giannotta, M., Trani, M., & Dejana, E. (2013). VE-cadherin and endothelial adherens junctions: Active guardians of vascular integrity. Developmental Cell, 26(5), 441–454. https://doi.org/10.1016/j.devcel.2013.08.020

17. Huynh, J., Nishimura, N., Rana, K., Peloquin, J. M., Califano, J. P., Montague, C. R., King, M. R., Schaffer, C. B., & Reinhart-King, C. A. (2011). Age-related intimal stiffening enhances endothelial permeability and leukocyte transmigration. Science Translational Medicine, 3(112). https://doi.org/10.1126/scitranslmed.3002761

18. Ito, S., Okuda, S., Abe, M., Fujimoto, M., Onuki, T., Nishimura, T., & Takeichi, M. (2017). Induced cortical tension restores functional junctions in adhesion-defective carcinoma cells. Nature Communications, 8(1), 1–16. https://doi.org/10.1038/s41467-017-01945-y

19. Juettner, V. V., Kruse, K., Dan, A., Vu, V. H., Khan, Y., Le, J., Leckband, D., Komarova, Y., & Malik, A. B. (2019). VE-PTP stabilizes VE-cadherin junctions and the endothelial barrier via a phosphatase-independent mechanism. Journal of Cell Biology, 23(11): 1024–1030. https://doi.org/10.1083/jcb.201807210

20. Komarova, Y. A., Kruse, K., Mehta, D., & Malik, A. B. (2017). Protein Interactions at Endothelial Junctions and Signaling Mechanisms Regulating Endothelial Permeability. Circulation Research, 120(1), 179–206. https://doi.org/10.1161/CIRCRESAHA.116.306534

21. Kuhne, W., Piper, H. M., & Diisseldorf, D.-. (1993). Disintegration of cytoskeletal in energy-depleted endothelial structure cells of actin filaments. https://doi.org/10.1152/ajpheart.1993.264.5.H1599

22. Le Duc, Q., Shi, Q., Blonk, I., Sonnenberg, A., Wang, N., Leckband, D., & De Rooij, J. (2010). Vinculin potentiates E-cadherin mechanosensing and is recruited to actin-anchored sites within adherens junctions in a myosin II-dependent manner. Journal of Cell Biology, 189(7), 1107–1115. https://doi.org/10.1083/jcb.201001149

23. Lorén, N., Hagman, J., Jonasson, J. K., Deschout, H., Bernin, D., Cella-Zanacchi, F., Diaspro, A., McNally, J. G., Ameloot, M., Smisdom, N., Nydén, M., Hermansson, A. M., Rudemo, M., & Braeckmans, K. (2015). Fluorescence recovery after photobleaching in material and life sciences: Putting theory into practice. Quarterly Reviews of Biophysics, 48(3), 323–387. https://doi.org/10.1017/S0033583515000013

24. Maruthamuthu, V., Sabass, B., Schwarz, U. S., & Gardel, M. L. (2011). Cell-ECM traction force modulates endogenous tension at cell-cell contacts. Proceedings of the National Academy of Sciences of the United States of America, 108(12), 4708–4713. https://doi.org/10.1073/pnas.1011123108

25. Miyake, Y., Inoue, N., Nishimura, K., Kinoshita, N., Hosoya, H., & Yonemura, S. (2006). Actomyosin tension is required for correct recruitment of adherens junction components and zonula occludens formation. Experimental Cell Research, 312(9), 1637–1650. https://doi.org/10.1016/j.yexcr.2006.01.031

26. Qin, D., Xia, Y., & Whitesides, G. M. (2010). Soft lithography for micro- and nanoscale patterning. Nature Protocols, 5(3), 491–502. https://doi.org/10.1038/nprot.2009.234

27. Salvi, A. M., & DeMali, K. A. (2018). Mechanisms linking mechanotransduction and cell metabolism. Current Opinion in Cell Biology, 54(Figure 1), 114–120. https://doi.org/10.1016/j.ceb.2018.05.004

28. Tabdili, H., Langer, M., Shi, Q., Poh, Y. C., Wang, N., & Leckband, D. (2012). Cadherin-dependent mechanotransduction depends on ligand identity but not affinity. Journal of Cell Science, 125(18), 4362–4371. https://doi.org/10.1242/jcs.105775

29. Tojo, K., Tamada, N., Nagamine, Y., Yazawa, T., Ota, S., & Goto, T. (2018). Enhancement of glycolysis by inhibition of oxygen-sensing prolyl hydroxylases protects alveolar epithelial cells from acute lung injury. FASEB Journal, 32(4), 2258–2268. https://doi.org/10.1096/fj.201700888R

30. Tzima, E., Irani-Tehrani, M., Kiosses, W. B., Dejana, E., Schultz, D. A., Engelhardt, B., Cao, G., DeLisser, H., & Schwartz, M. A. (2005). A mechanosensory complex that mediates the endothelial cell response to fluid shear stress. Nature, 437(7057), 426–431. https://doi.org/10.1038/nature03952

31. Wu, D., & Birukov, K. (2019). Endothelial cell mechano-metabolomic coupling to disease states in the lung microvasculature. Frontiers in Bioengineering and Biotechnology, 7(JUL), 1–18. https://doi.org/10.3389/fbioe.2019.00172

32. Brachman, P. S., & Abrutyn, E. (2009). Bacterial infections of humans: Epidemiology and control. In *Bacterial Infections of Humans: Epidemiology and Control*. Springer US. https://doi.org/10.1007/978-0-387-09843-2

33. Labrie SJ, Samson JE, Moineau S. Bacteriophage resistance mechanisms. Nat Rev Microbiol. 2010 May;8(5):317-27. 10.1038/nrmicro2315.

34. van Houte S, Ekroth AK, Broniewski JM, Chabas H, Ashby B, Bondy-Denomy J, Gandon S, Boots M, Paterson S, Buckling A, Westra ER. The diversity-generating benefits of a prokaryotic adaptive immune system. Nature. 2016 Apr 21;532(7599):385-8. doi: 10.1038/nature17436.

35. Mayer, A., Mora, T., Rivoire, O., & Walczak, A. M. (2016). Diversity of immune strategies explained by adaptation to pathogen statistics. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(31), 8630–8635. https://doi.org/10.1073/pnas.1600663113

36. Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. In *Cellular and Molecular Life Sciences* (Vol. 73, Issue 23, pp. 4433–4448). Birkhauser Verlag AG. https://doi.org/10.1007/s00018-016-2299-6

37. Pilosof, S., Alcalá-Corona, S. A., Wang, T., Kim, T., Maslov, S., Whitaker, R., & Pascual, M. (2020). The network structure and eco-evolutionary dynamics of CRISPR-induced immune diversification. *Nature Ecology and Evolution*, *4*(12), 1650–1660. https://doi.org/10.1038/s41559-020-01312-z

38. Sethna, J., Dahmen, K. & Myers, C. Crackling noise. *Nature* **410**, 242–250 (2001). https://doi.org/10.1038/35065675

39. Mehta, A. P., Mills, A. C., Dahmen, K. A., & Sethna, J. P. (2002). Universal pulse shape scaling function and exponents: Critical test for avalanche models applied to Barkhausen noise. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *65*(4), 6. https://doi.org/10.1103/PhysRevE.65.046139

40. Salje, E. K. H., & Dahmen, K. A. (2014). Crackling noise in disordered materials. *Annual Review of Condensed Matter Physics*, *5*(1), 233–254. https://doi.org/10.1146/annurev-conmatphys-031113-133838

41. Matthew C. Kuntz and James P. Sethna. Noise in disordered systems: The power spectrum and dynamic exponents in avalanche models. *Phys. Rev.* B **62**, 11699 – Published 1 November 2000 https://doi.org/10.1103/PhysRevB.62.11699

42. Childs, L. M., Held, N. L., Young, M. J., Whitaker, R. J., & Weitz, J. S. (2012). Multiscale model of crispr-induced coevolutionary dynamics: Diversification at the interface of Lamarck and Darwin. *Evolution*, *66*(7), 2015–2029. https://doi.org/10.1111/j.1558-5646.2012.01595.x

43. Childs, L. M., England, W. E., Young, M. J., Weitz, J. S., & Whitaker, R. J. (2014). CRISPR-induced distributed immunity in microbial populations. *PLoS ONE*, *9*(7). https://doi.org/10.1371/journal.pone.0101710

44. Kuntz, M (2001) Barkhausen Noise: Simulations, Experiments, Power Spectra, and Two Dimensional Scaling *Phys. Rev.* B **62**, 11699 https://doi.org/10.1103/PhysRevB.62.11699

45. William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. 2007. Numerical Recipes 3rd Edition: The Art of Scientific Computing (3rd ed.). Cambridge University Press

46. Alstott J, Bullmore E, Plenz D. Powerlaw: a Python package for analysis of heavy-tailed distributions. PLoS One. 2014 Jan 29;9(1):e85777. doi: 10.1371/journal.pone.0085777

47. A. Clauset, C.R. Shalizi, M.E.J. Newman, Http://Dx.Doi.Org/10.1137/070710111 51 (2009) 661–703.

48. Perkovic´, O., Perkovic´, P., Dahmen, K. A., & Sethna, J. P. (n.d.). *Disorder-induced critical phenomena in hysteresis: Numerical scaling in three and higher dimensions*.

49. Sheikh, M. A., Weaver, R. L., & Dahmen, K. A. (2016). Avalanche Statistics Identify Intrinsic Stellar Processes near Criticality in KIC 8462852. *Physical Review Letters*, *117*(26). https://doi.org/10.1103/PhysRevLett.117.261101

50. https://www.who.int/tools/flunet(https://app.powerbi.com/view?r=eyJrIjoiNjViM2Y4NjktMjJmMC00Y2NjLWFmOWQtODQ0NjZkNWM1YzNmIiwidCI6ImY2MTBjMGI3LWJkMjQtNGIzOS04MTBiLTNkYzI4MGFmYjU5MCIsImMiOjh9)

51. https://splunk-public.machinedata.illinois.edu/en-US/app/uofi_shield_public_APP/home(https://uofi.app.box.com/s/nrrnx2fn6sqt4hhuh2ckq2u6z5o4zmqg)

52. https://data.who.int/dashboards/covid19/data?n=c

53. William H. Press, Saul A. Teulkalsky, William T. Vetterling, Brian P. Flannery, *Numerical Recipies: The Art of Scientific Computing*, Third Edition (New York, NY: Cambridge University Press, 2007) p. 264-266

54. Uhl, J. T., Pathak, S., Schorlemmer, D., Liu, X., Swindeman, R., Brinkman, B. A. W., LeBlanc, M., Tsekenis, G., Friedman, N., Behringer, R., Denisov, D., Schall, P., Gu, X., Wright, W. J., Hufnagel, T., Jennings, A., Greer, J. R., Liaw, P. K., Becker, T., … Dahmen, K. A. (2015). Universal Quake Statistics: From Compressed Nanocrystals to Earthquakes. *Scientific Reports*, *5*. https://doi.org/10.1038/srep16493