Project 1: What can the Logistic Map explain about COVID dynamics?

Electronic turnin (code and pdf of your paper to UNM Learn) and PRINTED (in color if relevant) stapled, hard copy of your paper due **Thursday, Feb 17** at the start of class.

Your report should be 4-6 pages. The report will contain

- An Abstract that summarizes the purpose and main findings of your paper. Be sure to specifically describe your most interesting finding, i.e., if someone didn't read your paper, they should still understand the essence of what you found (not just the steps you took to find it).
- A Introduction that defines and contextualizes important terms and concepts
- A Methods & Results (combined) section that contains
    - Figures with captions that are self-explanatory (someone could look at the figure and the caption and understand it without reading the full paper)
    - Explanation of your Methods explaining how you generated each figure and any associated calculations with pointers to code and enough detail that someone could replicate your results.
    - A text description of results that summarizes each figure that explains and interprets that figure. Generally, you will need at least one paragraph per figure.
- A Discussion/Conclusions section that interprets and summarizes overall findings
- A References section that lists your citations
- A Contributions Statement that describes each team members' contributions
- Code turnin: You will turn in code and a pdf on Learn. Code can be a github link. Do NOT print out code.
- Follow the IEEE format for references (included in the LaTeX download). You can use Overleaf to collaboratively write your papers in LaTex. Check your grammar and spelling (Grammarly is free). Avoid long, run-on sentences. Make this paper easy to read. See resources to be posted on Learn for scientific paper writing. A grading rubric will also be posted.

Your goals for Project 1 are to 1) demonstrate understanding of the logistic map and information theoretic measures including mutual information and transfer entropy, 2) show that you understand the Walker paper and which of the claims are well-supported by additional analysis, and then 3) use the logistic map and concepts in the Walker paper to *attempt* to explain a feature of population dynamics of COVID-19 infections.

One relevant claim in Walker et al. (2012) is that "top-down causation dominates for collective states in regimes with $0.2 \leq$ epsilon ≤ 0.7" and that transfer entropy changes from primarily bottom-up to top-down for epsilon ≈ 0.2. Your analysis should further test this claim to extend analysis in the Detering and Wright paper.

Another claim is that many major evolutionary transitions might be marked by a transition in causal structure. They use logistic growth "as a toy model for demonstrating how such a transition can drive the emergence of collective behavior in replicative systems." The paper claims that the transition from bottom-up to top-down information flow drives the emergence of

collective behavior. How is this framework helpful (or not helpful) in understanding growth dynamics of COVID-19?

Rubric:
Writing clarity: Uses proper grammar and spelling in clear and easily understood text. No convoluted or run-on sentences. (10 points)

Abstract (5 points)

- State the purpose of the paper and motivate why it is important
- Briefly explain what methods you used
- **Summarize key results**
- Note: References are not needed in the abstract.

Introduction (15 points)

- First paragraph defines the logistic map, sensitive dependence on initial conditions, and transfer entropy including equations, citations for equations, and definitions of all variables in those equations. Explains the significance of each of these terms. (5 pts)
- Second paragraph that summarizes the purpose, methods and analysis of the Walker paper and how you evaluated the analysis in that paper in your own words. You can also allude to your conclusions about the validity of the claims in Walker. (5 pts)
- Third paragraph that explains your motivation for using the logistic map and the concepts in the Walker paper to explain the dynamics of population growth of COVID-19 infections. You should also briefly explain in what ways your approach was and/or was not successful. (5 pts)
- Include citations.

Methods & Results (55 points)

Your Results should center around the following figures, and methods should explain how you generated each figure. Each figure should be clear with easily readable and clearly labeled axes and legends (if required). Each figure should have a brief caption that summarizes the figure (and if needed, each subfigure) and specifies what variables are shown. Captions and figures should be self-contained and understandable without reading the whole paper.

Each figure and result should include a description of the method you used to produce the data or result, including citation of any software you used (ie., the JIDT tool).

Part 1 (12 points)
(Please use subheadings in the paper)

In Figure 1 show:
Figure 1a: Plot 4 time series generated by the logistic map (on one graph is fine). Show 2 values of R each with 2 slightly different initial conditions. You must choose different values of R

and different initial conditions than are shown in the Deterding and Wright paper. Plot population vs. time for 100 time steps as 4 lines that demonstrate chaotic and nonchaotic dynamics. Explain what values of R generate sensitive dependence on initial conditions and what features of your figure demonstrate it.

Fig 1b: Provide 4 Venn diagrams (4 scaled pairs of overlapping circles) that show for your data generated by the logistic map:

- The Shannon Entropy for the first n time steps and the last n time steps, where n is the number of time steps before the chaotic series diverges.
- The Mutual Information for the chaotic pair of populations (for the first n and the last n) and the periodic pair (for the first n and the last n).

Calculate the Transfer Entropy between two time series and explain the result.
The methods should explain how you discretized your data and ensured that you have enough data in each discrete bin to calculate probabilities for information metrics.

Part 2 (16 points)

Deterding and Wright provided additional evidence that TE switches from bottom up to top down between epsilon between approximately 0.2 and 0.3, but there is substantial uncertainty in both topdown and bottom up TE.

Fig 2a. Pick 3 values of epsilon in this range and calculate TE for $1 <= k <= 4$. Show error bars or shaded regions to indicate confidence intervals.

Fig 2b. Compare Mutual Information (MI) among subpopulations with a lag for some values of epsilon in this range. Improve upon the Walker figure by clearly indicating statistically significant results.

Your methods should describe your calculations. Your description of results should explain whether or not your analysis supports the assertions in Walker et all that “top-down causation dominates for collective states in regimes with $0.2 < \text{epsilon} < 0.7$” (particularly in the range $0.2 < \text{epsilon} < 0.3$). Is there a clear value of epsilon at which there is a reversal in the flow of information from bottom-up to top-down?

Part 3 (27 points)

Include figures and substantive analysis and interpretation (Part 3 should be a similar effort to Parts 1 and 2 combined).

Fig 3: Repeat Fig. 1b, but using real-world COVID time series. You can choose multiple countries or states, or whatever set of time series you would like to use. You may want to average, smooth, or transform the data before discretizing it. You should compare multiple sets of time series, and you may want to consider different time lags. Use your comparisons to test the following hypothesis: The dynamics of a particular COVID variant in one country (or region)

predict the later dynamics of that variant in other countries (or regions); this relationship is strongest if they have similar population demographics, vaccination rates and/or interventions in place (i.e. mask requirements or restrictions on gathering). Use MI and/or TE between time series to look for evidence that supports or refutes this hypothesis. You may focus on one aspect of this hypothesis (i.e. you might only consider vaccination percentages or population age). Explain why you might expect MI and TE between your chosen time series to be useful measures, and why you do or do not see evidence to support the hypothesis. Note: A thorough analysis that something does not exist is just as worthy as an analysis that it does. In either case, provide a thorough explanation.

Fig 4: Choose your own adventure: use the logistic map to investigate something else in the dynamics of COVID infections. You could consider how interventions affect carrying capacity or growth rates. You can extend the analysis from Part 3 to consider more countries, regions or factors. Or you might consider how the emergence of, and competition between, new variants changes dynamics; how interventions or transmissibility change dynamics in theory or practice; or how within host evolution and spread of virus (the lower level growth of virus population) affect the between host evolution and spread of virus (the higher level growth dynamics of the virus). For the latter point, it may be relevant that SARS-CoV-2 variants appear to emerge from evolution in immunocompromised patients and/or animal reservoirs so the (evolutionary) viral dynamics within individuals may affect the spread of the virus between individuals.

Be creative; have fun, do a careful analysis and explain your motivation and results clearly. Provide a clear figure (or two figures, but no more than three) to succinctly present your results. Explain your Methods and interpret your results.

Discussion & Conclusions (10 points)

- How effectively did the Walker paper demonstrate that the transition from bottom-up to top-down information flow drives the emergence of collective behavior?
- How well does the logistic map (modified and applied in whatever way you used it) explain some aspect of the population dynamics of people with COVID-19?
- What insights on the pandemic do you get from considering top-down and bottom-up drivers of COVID-19 dynamics?
- Provide 3 at least paragraphs. Refer to your figures and analysis above and the figures and analysis in the Walker paper.

References and Contributions statement (5 points)

- List of references cited in the text. Everyone should cite the Walker paper, the Mitchell textbook, perhaps the Flake book, the code and paper you were provided, and any downloaded code used to generate results. Use the references format shown in the IEEE template.
- A web link is sufficient to cite code.
- Failure to cite sources or code will result in a 0 for the Project.
- You may cite Wikipedia (Wikipedia is particularly informative about the logistic map), but do not quote large sections of Wikipedia--you must paraphrase and re-interpret any

source you cite in your own words. You may also cite videos for this class paper, but note that neither Wikipedia nor YouTube videos are appropriate to cite for publications

- You must cite all sources and all code that you use that you did not write yourselves. Failure to cite code or sources constitutes academic dishonesty and will result in a 0 on the assignment, a failing grade in the class, and a report to the University.

Contributions Statement: List the contributions of each team member. An example statement is: AB wrote the code that generated Figs. 1 and 3 and wrote the associated results paragraphs. XY wrote the code and results for Fig. 2. Both authors collaborated to write the code and results for Fig. 4. AB wrote the Introduction. XY wrote the Discussion and Conclusion. XY formatted the references and edited the final paper. AB & XY consulted with EF to understand how to implement Q and R, but AB & XY wrote their own original code. Note: the Contributions statement is required to receive a grade. If any particular team member(s) made an apparently unusually large or small contribution, the Professor will discuss this with the team to determine final grades.