

Research



**Cite this article:** Danon L, Lacasa L, Brooks-Pollock E. 2021 Household bubbles and COVID-19 transmission: insights from percolation theory. *Phil. Trans. R. Soc. B* **376**: 20200284.

<https://doi.org/10.1098/rstb.2020.0284>

Accepted: 12 April 2021

One contribution of 21 to a theme issue 'Modelling that shaped the early COVID-19 pandemic response in the UK'.

**Subject Areas:**

health and disease and epidemiology, behaviour, computational biology

**Keywords:**

networks, percolation theory, disease transmission, household bubbles

**Author for correspondence:**

Leon Danon

e-mail: [l.danon@bristol.ac.uk](mailto:l.danon@bristol.ac.uk)

<sup>†</sup>Members of JUNIPER (Joint UNiversities Pandemic and Epidemiological Research), a consortium of modelling groups from seven universities: Bristol, Cambridge, Exeter, Lancaster, Manchester, Oxford and Warwick. See <http://maths.org/juniper>.

# Household bubbles and COVID-19 transmission: insights from percolation theory

Leon Danon<sup>1,†</sup>, Lucas Lacasa<sup>2</sup> and Ellen Brooks-Pollock<sup>3,4,†</sup>

<sup>1</sup>Department of Engineering Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK

<sup>2</sup>School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK

<sup>3</sup>University of Bristol School of Veterinary Sciences, Langford BS40 5DU, UK

<sup>4</sup>NIHR HPRU in Behaviour Change and Evaluation, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 2BN, UK

LD, 0000-0002-7076-1871; EB-P, 0000-0002-5984-4932

In the era of social distancing to curb the spread of COVID-19, bubbling is the combining of two or more households to create an exclusive larger group. The impact of bubbling on COVID-19 transmission is challenging to quantify because of the complex social structures involved. We developed a network description of households in the UK, using the configuration model to link households. We explored the impact of bubbling scenarios by joining together households of various sizes. For each bubbling scenario, we calculated the percolation threshold, that is, the number of connections per individual required for a giant component to form, numerically and theoretically. We related the percolation threshold to the household reproduction number. We find that bubbling scenarios in which single-person households join with another household have a minimal impact on network connectivity and transmission potential. Ubiquitous scenarios where all households form a bubble are likely to lead to an extensive transmission that is hard to control. The impact of plausible scenarios, with variable uptake and heterogeneous bubble sizes, can be mitigated with reduced numbers of contacts outside the household. Bubbling of households comes at an increased risk of transmission; however, under certain circumstances risks can be modest and could be balanced by other changes in behaviours.

This article is part of the theme issue 'Modelling that shaped the early COVID-19 pandemic response in the UK'.

## 1. Introduction

Transmission of close contact infections, such as COVID-19, fundamentally depends on social interactions between individuals. Patterns of social contact determine the rate and extent of spread in a population. Social networks are complex and hierarchical owing to how society is structured. Social distancing has been one of the main methods for controlling SARS-CoV-2 transmission in the absence of a vaccine or effective pharmaceutical interventions. In the UK, social distancing measures have sought to limit social interactions outside households, owing to the inherent challenges of preventing household transmission. Limiting all non-household contacts should reduce the reproduction number to close to zero; however, in practice not all external contacts can be stopped.

Social distancing comes at a cost which disproportionately affects some individuals more than others. Detrimental effects include increased loneliness, social isolation, lack of physical and emotional support and reduced childcare provision, which can all be associated with a financial cost. The formation of 'bubbles', defined as small, non-overlapping, groups of households that are permitted to come into contact with each other, is intended to maintain the

benefits of social distancing while reducing the negative impacts of isolation. This, in essence, has the effect of creating one large household out of two or more smaller households. Social support bubbles have been used in various forms in countries including New Zealand, Belgium and the UK.

The precise definition of an allowed bubble has varied over time and between countries, making their effectiveness as a strategy difficult to quantify. The epidemiological impact of bubbles depends on rates of transmission and mortality, infrastructures for tracking, testing and isolating cases, protective work environments and safe transport modes for key workers. In the UK, support bubbles are bubbles of two households where one household contains a single adult or young children. Childcare bubbles are bubbles where a person from one household provides childcare for another household. Christmas bubbles are groups of three households who are permitted to mix closely between 23 December and 27 December 2020. In all cases, the bubbles must be exclusive and non-overlapping and contacts within a bubble should be treated as household contacts if someone in the bubble tests positive for COVID-19 [1].

In this work, we explore the impact that bubbling strategies have on the ongoing COVID-19 epidemic, to provide an evidence base and inform decisions. We use the distribution of household sizes in the UK to identify the UK conditions that will enable ‘bubbles’ to be safe and effective.

## 2. Methods

### (a) Percolation theory

Percolation theory has served as a direct analogy for infectious disease transmission. Originating in material science to describe the flow of liquid through a porous medium, percolation theory has developed through mathematical abstraction to map epidemic transmission models on a social network to the purely structural problem of percolation of an underlying network [2]. The epidemic threshold, above which an epidemic is highly likely, is deeply related to the concept of the percolation threshold, above which a percolating cluster appears with a high probability [3].

Whereas mathematical models have usually focused on network abstractions (lattices or random graphs), in the last 20 years, compelling empirical evidence has made it clear that the connection topology of real systems, such as household networks, is quite different from the idealized lattices or random graphs traditionally studied in percolation theory. Instead, realistic networks display stylized patterns midway between order and disorder [4], and understanding not only how epidemics spread [5] but also how percolation can take place in models of complex networks with realistic characteristics has been the subject of intense study (see [6,7] and references therein).

### (b) Constructing networks of households

To build the contact network of households, initially, we consider the realistic household size distribution from the UK 2011 census (table 1 and figure 1*a*) [8]. Then, we assume that all individuals in a household form a clique (a fully connected subgraph) and consider each household as a single node in the network, effectively going from a network of individuals to a network of households. To generate between-household connections, we assume that each individual in a household has the same probability of generating external connections, so that the degree of a household-node is proportional to its size (number of individuals). We say that a household with  $n$  individuals is of size  $n$  and

**Table 1.** Household size distribution.

size	count
1	7 067 261
2	7 998 031
3	3 641 569
4	3 031 078
5	1 085 188
6	386 784
7	92 734
8	35 292
9	14 799
10	6876

thus constitutes a single node with degree  $n$ , effectively making the assumption that each individual makes one external connection. Finally, as a baseline, we assume that individuals are connected to individuals in other households at random and build connections between household-nodes. This process of building the network of households is akin to a configuration model [9] (figure 1*b*). We call this the *baseline network* without household bubbles.

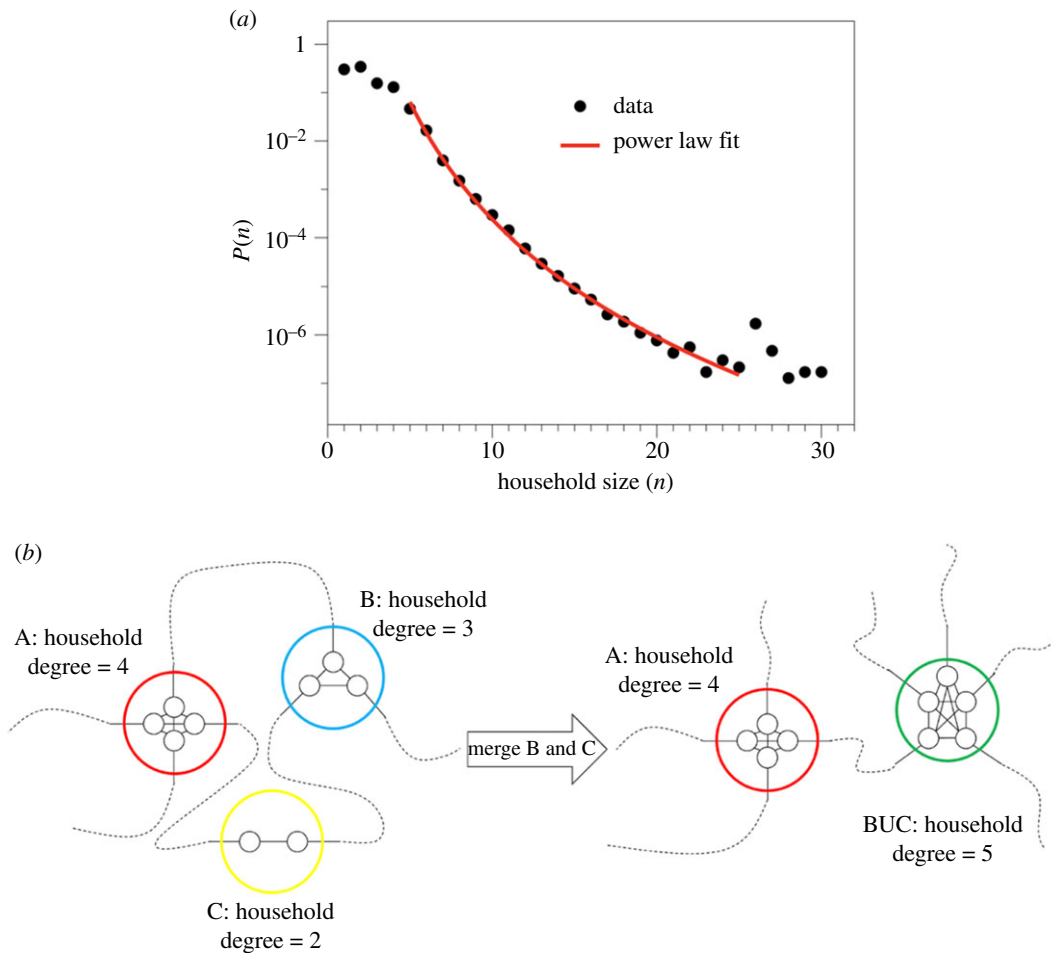
### (c) Simulating the effect of bubbling

We model bubbling by combining multiple households into larger ones, thus increasing the number of external links. Once the bubbled network has been formed, we study how resilient it is against random removal of these external links (see next section (§2*d*) on Percolation analysis), which will in turn allow us to associate a certain level of transmission risk to a specific bubbling scenario. We considered scenarios where all households behave in the same way, as well as more plausible scenarios with variable take-up and behaviours:

- **2-bubbles:** All households join together with one other household, chosen at random, to create a bubble of two households.
- **3-bubbles:** All households join together with two other households, chosen at random, to create a bubble of three households.
- **1 + 1:** All single-person households (size 1 households) join together with one other single-person household to make a two-person household.
- **1 +  $n$ :** All single-person households join together with another randomly chosen household of any size.
- **2 +  $n$ :** All households of size 1 or 2 join with another household of any size.
- **Plausible best case:** 33% of households form a bubble; half are 2-bubbles and half are 3-bubbles.
- **Plausible reasonable case:** 50% of households form a bubble; half are 2-bubbles and half are 3-bubbles.
- **Plausible worst case:** 75% of households form a bubble; half are 2-bubbles and half are 3-bubbles.

### (d) Percolation analysis

A component of a network is defined as a connected subgraph, i.e. as a set of nodes where each node is reachable to and from any other node following existing links. When the largest connected component of a network contains a finite, strictly positive fraction of the nodes of the network, we say the network has a giant component. Of course, all finite networks have a



**Figure 1.** Schematic of network construction, bubbling and percolation analysis. (a) The distribution of household sizes from the Office of National Statistics Census in 2011. (b) A schematic of the formation of a network at random and merging of households.

giant component (simply, the largest), but the concept reaches conceptual relevance in the limit of large networks.

In order to measure the relative robustness of networks before and after bubbling, we now perform a percolation analysis on the resulting networks, which essentially consists of measuring the proportion of links that need to be removed to break up the giant component. We begin by removing a proportion,  $(1-p)$ , of links from the network at random for all values of  $p$  from 1 to 0, decreasing in increments of 0.01. This is done independently at each value of  $p$ . As  $(1-p)$  is progressively increased, more between-household links are removed from the baseline network and the giant component decreases in size. At a critical point,  $p_c$  (the so-called percolation threshold) the network fragments abruptly and the giant component disappears. We use  $p_c$  as the indicative variable: a network with a low percolation threshold is much more risky with respect to a disease propagating over it, as only a handful of additional links are needed to get from the non-percolating to the percolating phase. For  $p$  ranging in  $[0,1]$ , we also measure the number of households in the giant component and the *average size* of the other components in the network (also known as the order parameter). As soon as the giant component emerges, it quickly accrues most of the nodes of the network, meaning that the other components will be smaller. A discontinuity (sharp peak) in the profile of this order parameter indicates the location of the phase transition, i.e. the threshold that distinguishes the non-percolating and percolating phases.<sup>1</sup> An alternative interpretation of this process is that we are generating a disease transmission network where transmission occurs with a probability  $p$  on a substrate of the baseline network and identifying the critical point in the disease transmission process.

We repeat this procedure for each network following different bubbling scenarios, estimating the location of the threshold for each case. Since each realization of the network is stochastic, we repeated this 100 times for each scenario and assessed the effects of intrinsic noise in the system.

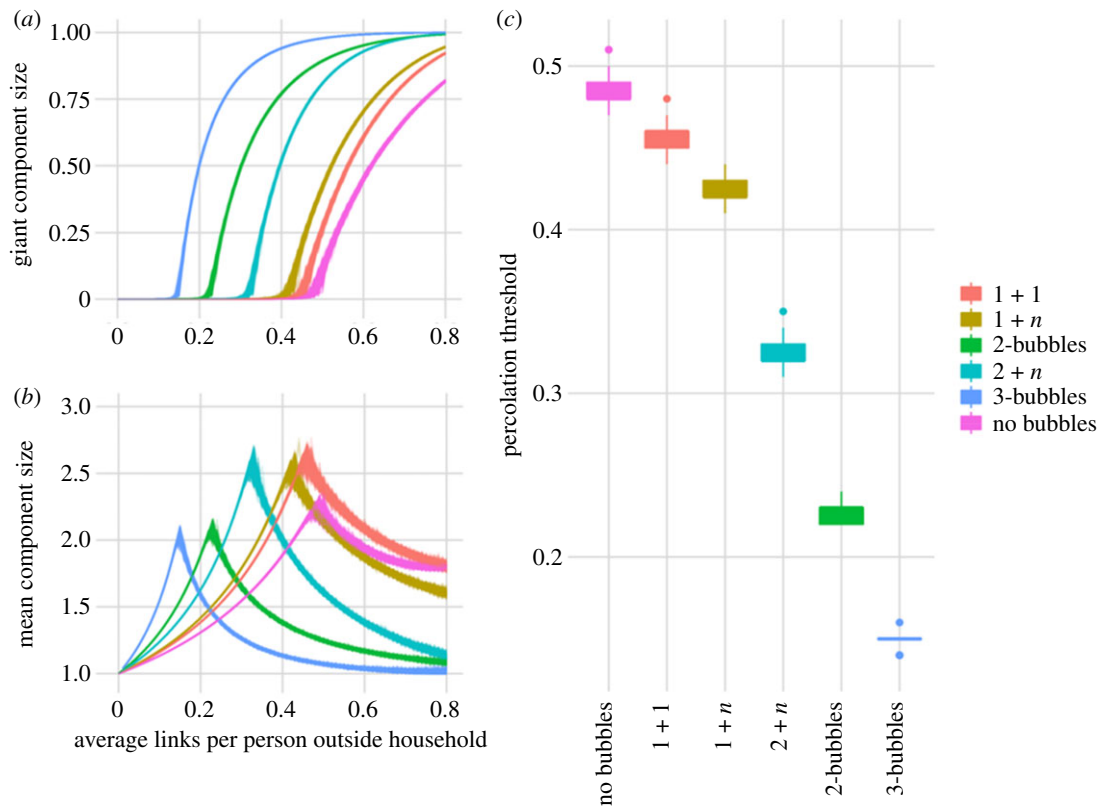
As described above, the relative location of this phase transition is used to infer the relative impact of specific bubbling strategies. For a lower percolation threshold, the underlying network is more resilient to link removal; in other words, for the same per-link probability of infection, transmission on a network with a low percolation threshold will affect a larger proportion of the population than a network with a high percolation threshold (figure 2). Moreover, if the system is at the critical point on the *baseline network*, the relative position of the threshold on equivalent networks under different bubbling assumptions will indicate whether a strategy is better (subcritical, no giant component) or worse (supercritical, large giant component).

### (e) Relating percolation thresholds to transmission potential

The percolation threshold is the point at which the giant component emerges (or disappears), and the network is globally connected (or disconnected). For locally tree-like networks,  $p_c$  can be expressed in terms of the degree and degree fluctuations of the network:

$$p_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle},$$

where  $\langle \rangle$  denotes the arithmetic mean.



**Figure 2.** Percolation analysis for hypothetical bubbling scenarios. (a) The proportion of households connected to the giant component. (b) The average size of components not connected to the giant component (order parameter) for the same bubbling strategies. (c) The percolation threshold for the different bubbling assumptions.

This purely topological property is in turn related to transmission potential and the reproduction number in the population [5,10].

Much of the focus in describing COVID-19 transmission has centred on the reproduction number, or  $R$ , defined as the average number of secondary cases caused by an average infected individual. However alternative reproduction numbers can be used to understand transmission at different scales, for example, the household reproduction number, which is the average number of secondary households infected by an average household [11–14].

The household reproduction number is given as

$$R_H = \sum_h \pi_h f_h R_G,$$

where  $f_h$  is the expected number of infections within a household of size  $h$ ,  $\pi_h$  is the proportion of households of size  $h$ , and  $R_G$  is the mean number of out-of-household infections due to a single infected person [11]. In our formulation,  $R_G = p$  and  $f_h$  is  $h$  multiplied by the secondary attack rate within households. We use the estimated secondary attack rate from UK data, which varies between 0.25 for large household sizes and 0.49 for households of size 2 [15,16]. Furthermore, below the percolation threshold,  $R_H < 1$  because the infection is necessarily limited to small, disconnected clusters.

All codes are available at doi:10.5281/zenodo.4659932 [17].

### 3. Results

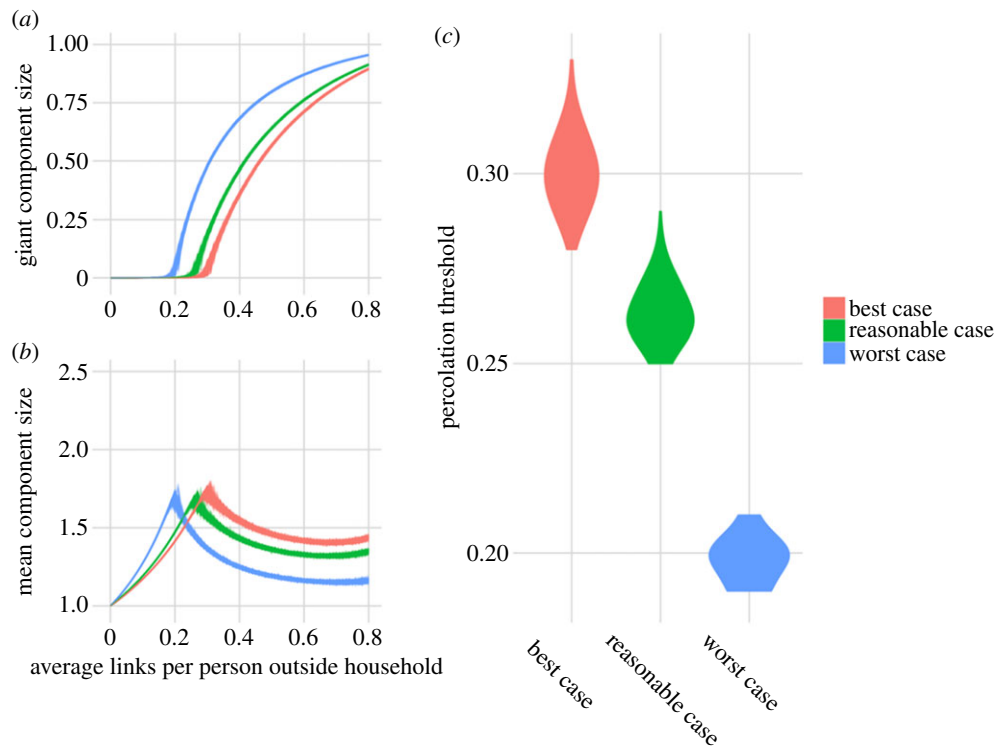
With no household bubbling, we estimate a percolation threshold of 0.48, meaning that, on average, if one out of every two people has a social contact outside of their household then a giant component emerges in the network of households

and there is the potential for large-scale outbreaks. At the percolation threshold, the household reproduction number is estimated to be 0.38, so even though a giant component exists, the epidemic is unlikely to affect a substantial proportion of the population since the reproduction number is less than 1. If each person in the household had a single link to other households, the reproduction number would be 0.79.

Single-person households (households of size 1) joining with another single-person household (scenario 1 + 1), or another household of random size (scenario 1 +  $n$ ) have a modest impact on network connectivity and transmission and the average bubble size increases by 0.4 and 0.5 person, respectively (figure 2). The percolation threshold is reduced by less than 15%. This translates to an increase in the household reproduction number of less than 0.3, and in practice this difference might not be observable with heterogeneity in household secondary attack rates (table 2).

In the scenario where all households form 2-bubbles, there is a substantial impact on network connectivity and transmission potential. The average bubble size increases to 4.7 people per bubble and the percolation threshold is decreased to 0.23 (figure 2), which means that one in four people with social contact outside their household is sufficient for the giant component to exist. Compared with no bubbling, 2-bubbles have the potential to increase network connectivity and increase the household reproduction number. For an average of one outside contact per person, the estimated household reproduction number is 1.38 (table 2).

All households forming 3-bubbles is the worst scenario we considered. As for 2-bubbles, the mean bubble size is increased, in this scenario to 7.1. The percolation threshold is 0.15, which means that a giant component forms if 1 in 6



**Figure 3.** Percolation analysis for plausible bubbling scenarios. (a) The proportion of households connected to the giant component. (b) The average size of components not connected to the giant component for the same bubbling strategies. (c) The percolation threshold for the different bubbling assumptions.

**Table 2.** Comparison of bubbling scenarios.

	average bubble size	theoretical $p_c$	measured $p_c$	$R_H$ when $p = 1$
no bubbles	2.4	0.476	0.48	0.79
2-bubbles	4.7	0.224	0.23	1.38
3-bubbles	7.1	0.146	0.15	1.81
1 + 1	2.8	0.448	0.44	1.11
1 + $n$	2.9	0.414	0.41	1.13
2 + $n$	3.9	0.319	0.33	1.27
best plausible case	2.9	0.290	0.30	0.92
reasonable plausible case	3.2	0.255	0.26	0.99
worst plausible case	4.3	0.193	0.19	1.23

people has an external social contact. For an average of one outside contact per person, the estimated household reproduction number is 1.81.

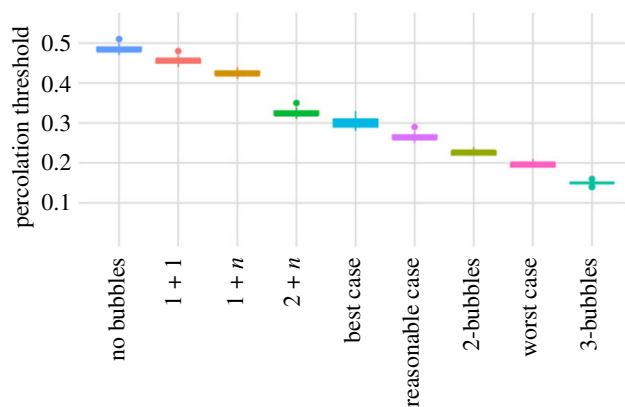
For plausible scenarios with variable uptake of bubbling and heterogeneity in bubble sizes, results are considerably less dramatic than when all households form a bubble. The best plausible case, with one-third of households forming 2-household and 3-household bubbles in approximately equal proportions, results in household-to-household transmission potential similar to  $1 + n$  bubbling. The reasonable plausible scenario, where 50% of households form a bubble of 2 or 3 households, results in network connectivity and transmission potential that is slightly greater than the best plausible case; however, the difference is relatively modest and might be unobservable for heterogeneity in secondary attack rates. The impact of reasonable plausible bubbling compared with no bubbling could be mitigated by individuals reducing their external

contact rate by 0.25, for example, a bubble of four individuals could reduce their external contacts from four to three.

Finally, the worst plausible scenario, where 75% of households form a bubble, produces similar characteristics to the 2-bubble scenario. Although the average bubble size is slightly smaller (4.3 for the worst plausible case versus 4.7 to 2-bubbles), the percolation threshold is lower for the worst plausible case owing to the occurrence of large bubbles that connect the network (figure 3 and table 1).

In figure 4, we show how the combined household bubbling strategies rank in terms of percolation threshold and therefore vulnerability to infection for all strategies considered. Small households joining up to each other or other households generally has a small effect, whereas larger households joining up with each other has a substantial effect. This remains the case when it is only a proportion of households that join up, as expected for Christmas bubbling scenarios.





**Figure 4.** The percolation threshold for all bubbling scenarios ordered from most vulnerable to link removal (left) to least vulnerable (right).

## 4. Discussion

In this analysis, we provide a simple network framework in which to quantify the effect of household bubbles on the transmission of COVID-19. We find that the creation of support bubbles between a single-person household and another household of any size has a small impact on transmission. The ubiquitous generation of bubbles for all households has the potential to make transmission extremely hard to control, therefore household bubbles should not be encouraged in general. However, we find that for intermediate uptake rates—for example where 50% of households form a temporary ‘Christmas’ bubble—that the additional transmission potential could be mitigated by a reduction in contacts outside the home.

There are natural parallels between percolation theory and disease dynamics, with analogies between the link probability and the giant component and the reproduction number and final epidemic size. Percolation theory has been used before to describe disease transmission in structured populations [5,10], although often using a theoretical framework, rather than to inform interventions (with the possible exception of [18]). With regard to the impact of bubbling, our results concord with Leng *et al.* [19], who concluded that bubbling of single-person households has a minimal impact on transmission [19].

There are a number of limitations to our analysis. First, we used a relatively simple network formulation in which the number of external contacts was proportional to household size. While this is true in general, in reality, there are more complex patterns where the number of external contacts

saturates with household size. Second, the model had no time dependence; therefore, we were not able to capture the formation and dissolution of bubbles, and this is particularly relevant for temporary festive bubbles. Third, the model contains no spatial component. It is likely that forming local bubbles is preferable to long-distance bubbles, but we were not able to investigate that question here. Fourth, we were limited by a lack of data on current bubbling practices. The Office for National Statistics opinions and lifestyle survey [20] reports that around 40% of adults in the UK have formed a support bubble in 2020, but we do not have data broken down by household size or data on future bubbling intentions. Our estimates of the household reproduction number are dependent on the secondary attack rate in households, which is also uncertain and probably varies with household size.

This work provides insight into the impact of bubbling on the transmission dynamics of COVID-19. We find that, in a UK setting, the formation of bubbles can be detrimental if taken up by a sizeable proportion of the population. Therefore, messaging around bubbling should be framed in a way that communicates the negative implications as well as the benefits. In particular, large gatherings of many households should be discouraged unless absolutely necessary. It is not clear how these conclusions would translate to other settings, where the definition of a household, the number of people per household and the out-of-household social mixing vary considerably. Future work could use data from other countries to explore the implications of bubbles for different social settings.

**Data accessibility.** Publicly available data from the Office for National Statistics.

**Authors’ contributions.** L.D. designed the study and performed the simulations. L.D., E.B.-P. and L.L. analysed the data and wrote the paper.

**Competing interests.** We declare we have no competing interests.

**Funding.** L.D. and E.B.-P. are supported by the Medical Research Council (MRC) (MC/PC/19067; MR/038613/1). L.L. acknowledges the financial support of the EPSRC via Early Career Fellowship EP/P01660X/1. E.B.-P. is partly funded by the NIHR Health Protection Research Unit (HPRU) in Behavioural Science and Evaluation.

**Acknowledgements.** We thank the members of SPI-M for useful suggestions and comments and Joseph Hilton for providing the household data in a convenient manner.

## Endnote

<sup>1</sup>Indeed, in the thermodynamic limit, when the number of nodes becomes very large, the order parameter diverges [3].

## References

1. Department of Health and Social Care. 2020 *Making a support bubble with another household*. See <https://www.gov.uk/guidance/making-a-support-bubble-with-another-household>.
2. Broadbent SR, Hammersley JM. 1957 *Percolation processes. I. Crystals and mazes*. *Math. Proc. Camb. Phil. Soc.* **53**, 629–641. (doi:10.1017/S0305004100032680)
3. Stauffer D, Aharony A. 1992 *Introduction to percolation theory*. London, UK: Taylor & Francis.
4. Latora V, Nicosia V, Russo G. 2017 *Complex networks: principles, methods and applications*. Cambridge, UK: Cambridge University Press.
5. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. 2015 Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979. (doi:10.1103/RevModPhys.87.925)
6. D’Souza RM, Nagler J. 2015 Anomalous critical and supercritical phenomena in explosive percolation. *Nat. Phys.* **11**, 531–538. (doi:10.1038/nphys3378)
7. Lee D, Kahng B, Cho YS, Goh K-I, Lee D-S. 2018 Recent advances of percolation theory in complex networks. *J. Korean Phys. Soc.* **73**, 152–164. (doi:10.3938/jkps.73.152)
8. Office for National Statistics. 2011 *2011 Census: Population and household estimates for the United Kingdom, March 2011*. See <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationandhouseholdestimatesfortheunitedkingdom/2011-03-21#number-of-people-in-households>.
9. Molloy M, Reed B. 1995 A critical point for random graphs with a given degree sequence. *Random Struct. Algorithms* **6**, 161–180. (doi:10.1002/rsa.3240060204/abstract)

10. Newman MEJ. 2002 Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128. (doi:10.1103/PhysRevE.66.016128)
11. Goldstein E, Paur K, Fraser C, Kenah E, Wallinga J, Lipsitch M. 2009 Reproductive numbers, epidemic spread and control in a community of households. *Math. Biosci.* **221**, 11–25. (doi:10.1016/j.mbs.2009.06.002)
12. Ball F, Sirl D, Trapman P. 2010 Analysis of a stochastic SIR epidemic on a random network incorporating household structure. *Math. Biosci.* **224**, 53–73. (doi:10.1016/j.mbs.2009.12.003)
13. Pellis L, Ball F, Trapman P. 2012 Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of  $R_0$ . *Math. Biosci.* **235**, 85–97. (doi:10.1016/j.mbs.2011.10.009)
14. Ball F, Pellis L, Trapman P. 2016 Reproduction numbers for epidemic models with households and other social structures II: comparisons and implications for vaccination. *Math. Biosci.* **274**, 108–139. (doi:10.1016/j.mbs.2016.01.006)
15. Ladhani SN *et al.* 2020 Secondary attack rate and family clustering of SARS-CoV-2 infection in children of healthcare workers with confirmed COVID-19. *Clin. Infect. Dis.* **2020**, ciaa1737. (doi:10.1093/cid/ciaa1737)
16. Bernal JL *et al.* 2020 Transmission dynamics of COVID-19 in household and community settings in the United Kingdom. *medRxiv*, 2020.08.19.20177188. (doi:10.1101/2020.08.19.20177188)
17. Danon L, Lacasa L, Brooks-Pollock E. 2021 *Idanon/Percolation: household bubbling code. Apr 2*. (doi:10.5281/zenodo.4659933)
18. Davis S, Trapman P, Leirs H, Begon M, Heesterbeek JAP. 2008 The abundance threshold for plague as a critical percolation phenomenon. *Nature* **454**, 634–637. (doi:10.1038/nature07053)
19. Leng T *et al.* 2020 The effectiveness of social bubbles as part of a Covid-19 lockdown exit strategy, a modelling study. *Wellcome Open Res.* **5**, 213. (doi:10.12688/wellcomeopenres.16164.1)
20. Office for National Statistics. 2020 *Coronavirus and the social impacts on Great Britain: 21 August 2020*. See <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandwellbeing/bulletins/coronavirusandthesocialimpactsongreatbritain/21august2020>.