Virus testing optimisation using Hadamard pooling

Godfrey S. Beddard(1,2)* & Briony A. Yorke(2)*

(1) School of Chemistry, University of Edinburgh, David Brewster Road, EH9 3FJ (2) School of Chemistry, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, United Kingdom

* Corresponding Authors: G.S.Beddard@leeds.ac.uk, B.A.Yorke@leeds.ac.uk

Abstract

Pooled testing is an established strategy for efficient surveillance testing of infectious diseases with low-prevalence. Pooled testing works by combining clinical samples from multiple individuals into one test, where a negative result indicates the whole pool is disease free and a positive result indicates that individual testing is needed. Here we present a straightforward and simple method for pooled testing that uses the properties of Hadamard matrices to design optimal pooling strategies. We show that this method can be used to efficiently identify positive specimens in large sample sizes by simple pattern matching, without the requirement of complex algorithms.

Background

Many countries have used the 'test, trace and isolate' strategy to control the spread of highly infectious diseases such as SARS-CoV-2. The global capacity for clinical testing is limited by factors including cost, availability of reagents and testing capacity of clinical laboratories. The gold-standard for detecting SARS-CoV-2 is RT-PCR [1], this technique is expensive, requires specialised equipment and reagents, and is relatively slow, taking a few hours per test on average [2, 3].

During the height of the COVID-19 pandemic lateral flow testing (LFT) provided a solution to issues regarding laboratory testing capacity but were not without limitations. The 'test, trace and isolate strategy' relies on accurate, reproducible testing and data monitoring, however, LFT self-testing is hindered by test variability, false negatives, and data loss due to errors in self-reporting [4]. In combination, SARS-CoV-2 testing resulted in tens of thousands of tons of plastic waste [5], and unsustainable consumption of reagents [6] and gold nanoparticles [7].

Strategies to address this issue have resulted in the development of biodegradable and recyclable

biosensors [8, 9], methods to recover gold nanoparticles [10] and microfluidic and reagent free testing devices [11–14]. Despite these developments, individual testing remains inefficient when the fraction of positives in a population is small, i.e. when almost all tests are expected to be negative. However, this is just the situation that needs surveillance-type testing, so that a new pandemic can be anticipated. When the infection rate is lower and urgency is not so pressing, one way of increasing capacity is to group or pool samples. This batch analysis was first suggested by Dorfman [15] more than 70 years ago, but has received renewed interest due to the critical testing bottlenecks that were highlighted during the COVID-19 pandemic [16–18].

Sample Pooling

Dorfman's method consists of two stages, the first is to combine samples from multiple individuals for analysis. If the test result is negative then all samples are negative, if the test is positive for the pool then each sample is tested again individually. The efficiency of this method depends on the positivity rate and can be improved by applying various strategies generally defined as adaptive, non-adaptive and hybrid [19]. Non-adaptive pooling is performed by generating a number of pools according to a predefined combinatorial design, all of the pools are testing in parallel before identifying the positive sample by deconvolution with an algorithm based on the combinatorial design [20–23]. Adaptive strategies are performed in series, data concerning transmission and the results of each test informs which samples will be included in the next test [24–26] . Hybrid methods involve multiple rounds of combinatorial pooling and all samples are tested in parallel during each round [27, 28].

We propose a new method, Hadamard pooling, which is suitable for single round testing when the rate of positivity is low, or which can be used in a hybrid approach when the positivity rate is higher. Hadamard pooling is based upon the orthogonality properties of Hadamard **S** matrices that can identify an individual positive sample in a pool. This method is also compatible with multiplex testing, in which multiple targets can be identified in a single assay, for example by RT-PCR [29, 30], CRISPR-Cas9 assays [31] or colourimetric RT-LAMP analysis [32]. Hadamard pooling has the potential to increase the accuracy and efficiency of surveillance testing programmes to track the emergence and spread of infectious diseases in the event of a new pandemic.

Hadamard Pooling

Hadamard **S** matrices (Figure 1a) have found application in numerous instances of experimental optimal design including signal processing [33–36], imaging [8, 37, 38], X-ray crystallography [39] and spectroscopy [40–45]. In each of these types of experiments, a signal is modulated by the pattern of

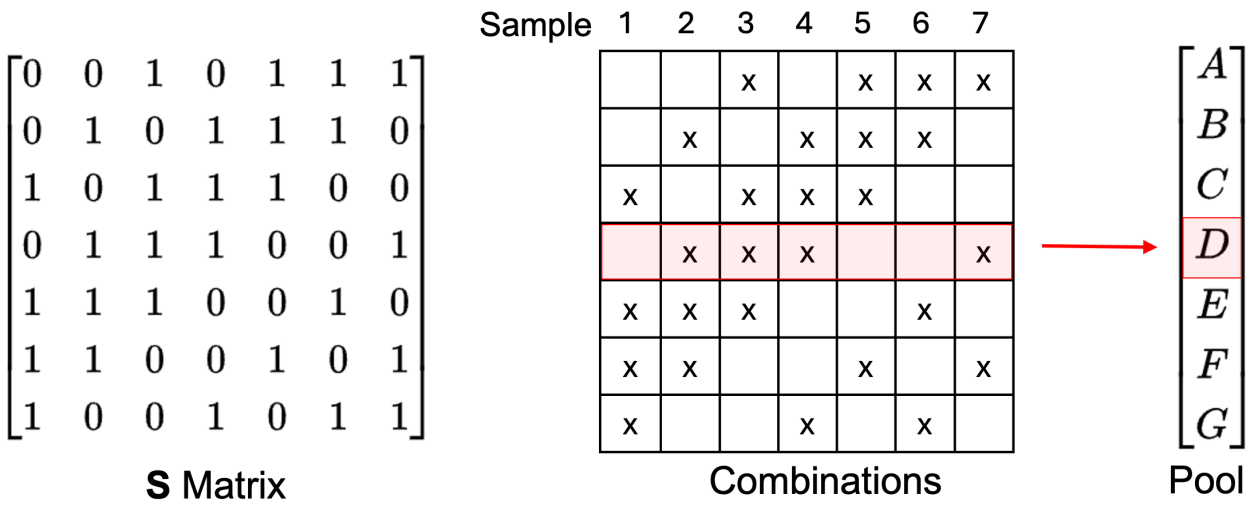

Figure 1: Hadamard pooling strategy scheme based on a 7×7 Hadamard matrix. Pools A - G are generated by grouping together 4 of the 7 individual samples according to the sequence of 1's and 0's in each row of the matrix where 1 indicates a sample should be added to the pool and 0 indicates that it is omitted. In the case of pool A samples 1, 2, 3 and 5 would be grouped.

a row in the matrix and summed on the detector. After this a linear transform with the inverted Hadamard **S** matrix returns the individual data. By summing up the signal its size is increased, but random noise of detection is averaged towards a constant value and therefore the signal-to-noise can be improved. The process of using the matrices and different ways of generating their patterns is described in supporting information (S1 appendix) and in references [39, 42, 46], Hadamard matrices in general are described in reference [46].

In the context of pooling methods, individual samples can be grouped into multiple pools according to each row of an **S** matrix with equivalent order to the number of samples being tested. Figure 1 shows a 7×7 **S** matrix to illustrate the pooling approach, in practice one of many larger Hadamard matrices could be used, although Hadamard matrices can only be generated with specific sizes, they are numerous.

In the example shown in Figure 1, n tests are needed for n samples and so there is no improvement over efficiency in comparison to testing individuals. However, in practice only a fraction of the matrix is needed to successfully identify a positive individual. For the purpose of demonstration a set of seven samples in which one is positive are used. Three pools A, D and F are generated by grouping four of the samples together according to a reduced 3×7 matrix constructed from the corresponding three rows of the 7×7 **S** matrix (Figure 2). Since the pattern of each column in the matrix is unique it is possible to identify which particular individual is positive. This is done by matching the pattern of results from a specific test, to the corresponding column of the reduced matrix. In the example shown in Figure 2, pools A and F return a negative test result, while pool D is positive. The results from

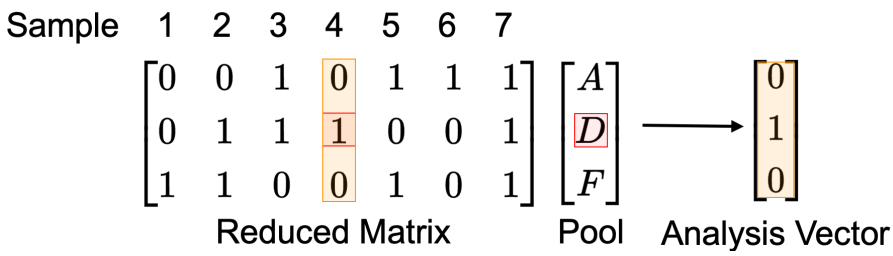

Figure 2: Reduced matrix with samples numbered 1 to 7 grouped into pools A, D and F. Red indicates which sample and pool is positive. The analysis vector elements indicate the results of the pool testing where 1 indicates a positive analysis result and 0 a negative one. Orange highlights the column of the reduced matrix that matches the analysis vector and therefore indicates that sample 4 is positive.

each pool are recorded in a vector where 0 indicates a negative pool and 1 indicates a positive pool. This analysis vector is then compared to each column in the reduced matrix, if the analysis vector and column match then this indicates that the corresponding sample is positive. For the case shown in Figure 2 the analysis vector, [0 1 0], matches the column in the reduced matrix corresponding to sample 4.

With this method if, for example, eleven specimens need to be checked the nature of the 11×11 Hadamard matrix is such that only four measurements are needed to confirm a result provided, that no more than one specimen is positive. The pattern of the 11×11 matrix is shown, with others, in the Supporting information (S1 appendix). Similarly, assuming that just one sample is positive, only four rows of the $n = 15$ and six of the $n = 31$ Hadamard $\mathbf{S}$ matrices are needed, Table 1.

The condition of low prevalence and hence a low probability of a positive test should be the general expectation when large swathes of a population are tested and the infection is not spreading exponentially i.e. reproduction ratio $R < 1$. The case when more than one sample is positive is examined next as this involves choosing rows more carefully.

Choosing Rows

If, for example, seven samples are tested using an $\mathbf{S}$ matrix size of, say 15×15, then combinatorially a large number of rows are available to determine which pools to generate, e.g., $\frac{15!}{7!(15-7)!} = 6435$ and this increases very rapidly with the size of the matrix. Combinations of rows with columns that contain the same pattern of ones and zeros must be excluded. For the example in Figure 1, using pools A, B and D is not suitable because columns 5 and 6 will have the same pattern in the corresponding reduced matrix and therefore the analysis vector cannot distinguish whether sample 5 or 6 is positive. A further condition is imposed when the number of positive samples is either one or two. Rows must be selected such that when two of their columns are added together they are not equal to a third. It is also assumed that a sample's result is not quantitative, and that the measurement indicates only

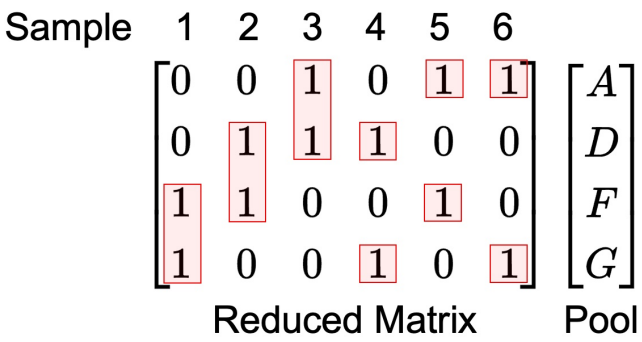

Figure 3: The optimal reduced matrix based on a 7×7 Hadamard $\mathbf{S}$ matrix with corresponding pools.

a binary result, i.e. either positive or negative, thus 1 or 0 and that two positive samples in the same row are measured as 1 in total and not 2. The consequence of this latter restriction is that one column of the $\mathbf{S}$ matrix must be removed (reducing the total number of samples) and in the case of the 7×7 matrix one row is added to the reduced matrix, meaning that four pools are now needed from 6 samples. The only acceptable combination of pools for this matrix is A, D, F, G, shown in Figure 3.

For larger matrices there are several acceptable groups of rows but these are still a very small fraction, typically $< 1\%$, of the total number of combinations. In the 11×11 matrix five rows are needed, for example B, D, E, F, J and with this combination of rows the last column in the matrix must be deleted. Different combinations of rows are generated if an alternative column is deleted. Suitable rows for other matrices are listed in Table 1. The minimum number of rows required appears to be about half for a small matrix and slightly less for larger ones [9].

Results and Discussion

This method is suitable when the likelihood of measuring two or more positive samples is low. Assuming a Poisson (or Binomial) distribution, with a positivity rate of 1%, and using a 7×7 matrix one positive sample has a 30% of being present in the whole matrix or 6.5% chance of being present in a column but two or more positives have a 9% in total or 0.23% chance in a column. The same numbers for columns in a 15×15 matrix are 13% and 1% respectively. These results were confirmed by simulating the process stochastically. Typically 20,000 repeat calculations were averaged and the simulations closely match the predicted behaviour from the Poisson (and Binomial) distribution (Figure 4).

When two positive samples are present, the result produced does not correspond to any column and their identities, within strict limits, can be estimated but usually re-testing would be required. For example in Figure 2, two positive pools returning analysis vector [0 1 1 1] could be made up of samples $1+2, 1+4, 1+5, 1+6$ so that only sample 1 may be positively identified. This is case (ii)

n	Row labels. 1 positive case (i)	Row labels, >=1 positive , case (ii)	Row labels, >=1 positive case (iii)	1 % positive. Chance as % of 0, 1, $\geq$ 2 present
7	A D F	A D F G	——	93.2, 6.53, 0.23
11	B D E F	B D E F J	A C G H I K	89.6, 9.85, 0.56
15	B C D E	B C D E G I J	A B C D E F G H M N	86.1, 12.9, 1.02
19	B E F H N	B E F H J K L Q R	A D G I K M N O P S	82.7, 15.7, 1.59
23	A H K N Q	A H K L O P R T V W	A F H I K O P R T V W	79.5, 18.3, 2.27
27	A B E N S	D G I J L N P S U V Z	A B C D E F G H J M O T X	76.3, 20.6, 3.05
31	B C E G L N	B C E G H I J L Q S U c	C D G J L N P U V X Y a e	73.3, 22.7, 3.92

Table 1: Suitable pool combinations when using Hadamard **S** matrices initially of size $n \times n$. The last column was ignored in choosing the **S** matrix for each list of rows in cases(ii) and (iii). In the larger matrices there are equivalent rows that could be used. The right-hand column shows the probability as a percentage of detecting $0, 1$ or ≥ 2 positives, when there are 1% positive samples out of n, calculated using the Poisson distribution. The letters follow the order 'A, B, C, . . Z, a, b, c, . . z'.

in Table 1. If a quantitative method of analysis is used then identifying the samples is easier as these same pairs each have a unique analysis vector, for example [0 1 2 1] for the pair 1 & 2. The only case needing re-testing is when the analysis vector is [1 1 1 1] for which there are three combinations out of a possible 15 that are not unique; 1+3, 2+6 and 4+5. To overcome this restriction another row(s) could be used and such a pattern is shown in case (iii) in Table 1 with the exception of $n = 7$ where such a row does not exist. The 15×15 matrix is slightly unfavourable in this respect, whilst the others are more efficient needing only one extra row.

Large sample sizes

For a large number of samples with low positivity rate, a strategy similar to that of Dorfman [15] can be used. In the Dorfman method, all the specimens are pooled into a few large groups and then each of those that tested positive is retested. Our approach is similar but differs slightly in that these initial groups are tested according to the Hadamard method and then the specimens comprising any positive ones are tested again, also using the Hadamard method. Thus for 225 samples (15×15 matrix) and if only one sample is positive the Dorfman method involves testing 15 pooled groups then 15 more tests of the group containing the positive one. Using the Hadamard method, if only one positive is known

to be present and 4 rows (Table 1) are used to generate the pools, then four further measurements are needed from the group that contains the positive result, or $2m$ where m is the number of rows used. This is 3.6% of the total sample when $n = 15$ or a test/person ratio of 0.036. These values are effectively the same as the 0.1% positives shown as blue circles in Figure 4.

In the general case the number of positive samples should be present in proportions given by the Poisson distribution. If more than one positive is found in the first round, each row is individually summed and retested, to find all positives. The positive groups are then tested to give a total of $4+15+4+4$ tests in the 15×15 matrix which is $3m+n$ in general. If m rows are used from a matrix of size n and $i = 0, 1, 2$ positive samples occur with the fraction f_i then the number of tests per person T is the weighted sum of zero, one, two or three samples etc.,

$$T \approx \frac{1}{n^2}(mf_0 + 2mf_1 + (3m+n)f_2 + (4m+n)f_3 + \cdots)) \tag{1}$$

A plot of T is shown in Figure. 4 with data points taken from the literature and the Dorfman function for the number of tests per person:

$$T_D = 1 + \frac{1}{w} - (1-p)^w \tag{2}$$

Where w is the sample size and p the probability of infection ($p = 0.01$) with these conditions the minimum sample size is:

$$w_{min} = 1/\sqrt{p} \tag{3}$$

The number of tests required per person subsequently increases with the sample size, in contrast using the Hadamard method the number of tests decreases with increasing sample size. Additionally, the number of tests is significantly lower than that of the Dorfman method, and comparable to [16] for large numbers of samples. The Hadamard pooling method produced similar results to previous pooling strategies, but has the advantage that no complicated algorithm is needed to interpret the results, only simple pattern matching. The strategy of using the minimum number of tests m initially (case (i) Table 1) and then testing all n samples only when more than one is positive, and then as necessary reverting to testing only m rows, turns out to be more efficient than initially testing the full matrix first. Instead of testing all n samples in the second step the reduced matrix could be used, case (iii) Table 1, and would lead to a slightly better test/person ratio than shown in Figure 4 for 1% positives. If the full matrix has to be evaluated the simplest way to identify a positive sample using the Hadamard method is to left-multiply the column of results by the inverse of the **S** matrix, see Supplementary Information (S1 appendix). In the 1% positive case only the smaller matrices are

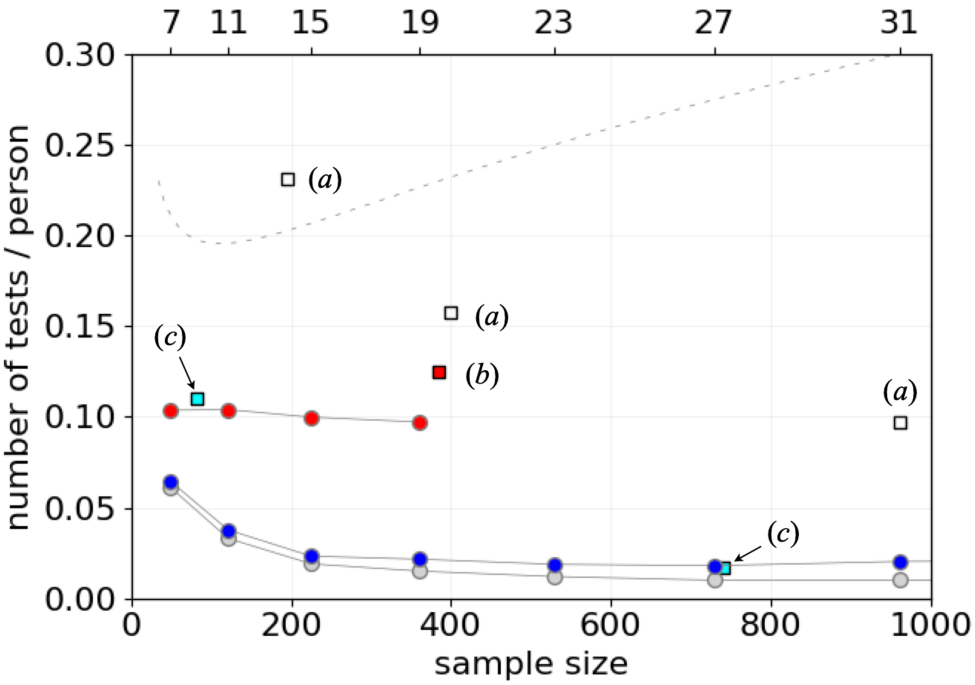

Figure 4: Figure 4. The number of tests / person vs. the sample size (matrix size squared, n^2) in a two step or Dorfman consecutive (two-stage or hierarchical) analysis with 0.1 and 1% positive samples. The top axis shows the **S** matrix size m. The blue circles are the data when 0.1 % on average are positive and the red circles when 1% are positive and therefore 0, 1, 2 or 3 etc. positives are present but weighted by the Poisson distribution and calculated using Equation 1. The grey circles are a stochastic simulation of the process when 0.1% are positive using rows as in case (i) Table 1. The similar calculation for 1% positives follows the red circles. Points labelled (a) are taken from Ghosh et al. [22], point (b) from Shental et al. [47], and (c) at 1 and 0.1 % positives from Figure 3 in Mutesa et al [16]. The dashed line is the Dorfman function.

included as multiple positives in each row or column soon occur as the matrix size increases and this involves using an increasingly involved sequence of testing for no gain in the number of tests/person.

The pooling method can be extended further for larger numbers, for example with three stages, $15^3 = 3375$ samples, between 12 and 21 measurements are needed. However, this and larger numbers are probably only rarely going to within the capacity of a single testing station and, although dilution of samples is possible [16, 26], such huge dilution may not be desirable or even feasible in practice.

Conclusion

We have demonstrated a straightforward approach to pooled testing using Hadamard **S** matrices to design pooling strategies. The efficiency of Hadamard pooling increases with sample size making the approach suitable for surveillance testing in low prevalence settings, ideally when the positivity rate is $\leq 1\%$. Hadamard pooling not only provides improvements in experimental screening efficiency but also in computational efficiency. A positive sample can be identified from a single round of testing by pattern matching, removing the need for computationally expensive decoding algorithms. This

strategy has the potential to contribute to the sustainable and efficient screening programmes needed to ensure preparedness for the next pandemic.

Methods

Matrix Generation

All calculations were performed using Python 3 (v 3.9) with the NumPy package (v 1.23) within a Jupyter Notebook (v 6.4). Hadamard matrices were generated using the Quadratic Residue, Shift Register or Doubling methods, as appropriate, and described in [33, 39, 46, 48].

To find suitable combinations of the Hadamard matrix rows as shown in Table 1, the combinations $C_{r,n} = n!/(r!(n-r)!)$ of r rows out of a total of n was calculated. For example selecting 10 rows of the $n = 15 \times 15$ matrix produces 3003 different combinations, but this rises rapidly to over 200 million for the $n = 31$ matrix with 13 rows. For smaller numbers of combinations each was exhaustively checked. The values of r were increased until a satisfactory list was found, for example a list with the conditions of case(iii), Table 1. There is only one smallest list for a $n = 11 \times 11$ matrix, i.e. ACGHIK and 65 smallest but equivalent lists out of 92378 when $r = 10, n = 19$ one of which is shown in Table 1. For larger matrices there are also several acceptable lists, but still a small fraction of the total (<1%), and only one of them is given in Table 1. (The exception is $n = 15$ where there are many (≈ 13 %) suitable combinations). When the number of combinations runs into millions, as it does for larger matrices, an exhaustive search is not feasible so the combinations were chosen at random until a suitable one was found. This is possible because, luckily, it happens that the suitable combinations tend to cluster and so can be accessed more quickly by random selection than a linear search.

Pooling Simulation

The person/test ratio was calculated for different sample sizes using Equation 2 assuming m rows used out of an $n \times n$ matrix with values $f_0, f_1 \cdots$ calculated from the Poisson distribution for the percentages as shown in the Figures 2 and 3.

References

[1] K. Munne, V. Bhanothu, V. Bhor, V. Patel, S. D. Mahale, and S. Pande, "Detection of sars-cov-2 infection by rt-pcr test: factors influencing interpretation of results," *Virusdisease*, vol. 32, no. 2, pp. 187–189, 2021.

[2] L. Cheng, L. Lan, M. Ramalingam, J. He, Y. Yang, M. Gao, and Z. Shi, "A review of current effective covid-19 testing methods and quality control," *Archives of Microbiology*, vol. 205, no. 6, p. 239, 2023.

[3] J. Budd, B. S. Miller, N. E. Weckman, D. Cherkaoui, D. Huang, A. T. Decruz, N. Fongwen, G.-R. Han, M. Broto, C. S. Estcourt, *et al.*, "Lateral flow test engineering and lessons learned from covid-19," *Nature Reviews Bioengineering*, vol. 1, no. 1, pp. 13–31, 2023.

[4] M. N. Esbin, O. N. Whitney, S. Chong, A. Maurer, X. Darzacq, and R. Tjian, "Overcoming the bottleneck to widespread testing: a rapid review of nucleic acid testing approaches for covid-19 detection," *Rna*, vol. 26, no. 7, pp. 771–783, 2020.

[5] Y. Peng, P. Wu, A. T. Schartup, and Y. Zhang, "Plastic waste release caused by covid-19 and its fate in the global ocean," *Proceedings of the National Academy of Sciences*, vol. 118, no. 47, p. e2111530118, 2021.

[6] R. T. Wulff, Y. Qiu, C. Wu, D. P. Calfee, H. K. Singh, I. Hatch, P. A. Steel, J. E. Scofi, L. F. Westblade, and M. M. Cushing, "Laboratory interventions to eliminate unnecessary rapid covid-19 testing during a reagent shortage," *American Journal of Clinical Pathology*, vol. 158, no. 3, pp. 401–408, 2022.

[7] F. Araste, A. D. Bakker, and B. Zandieh-Doulabi, "Potential and risks of nanotechnology applications in covid-19-related strategies for pandemic control," *Journal of Nanoparticle Research*, vol. 25, no. 11, p. 229, 2023.

[8] C. Hwang, N. Park, E. S. Kim, M. Kim, S. D. Kim, S. Park, N. Y. Kim, and J. H. Kim, "Ultrafast and recyclable dna biosensor for point-of-care detection of sars-cov-2 (covid-19)," *Biosensors and Bioelectronics*, vol. 185, p. 113177, 2021.

[9] T. Švarc, P. Majerič, D. Feizpour, Jelen, M. Zadravec, T. Gomboc, and R. Rudolf, "Recovery study of gold nanoparticle markers from lateral flow immunoassays," *Materials*, vol. 16, p. 5770, 2023.

[10] A. Sun, P. Vopařilová, X. Liu, *et al.*, "An integrated microfluidic platform for nucleic acid testing," *Microsyst Nanoeng*, vol. 10, p. 66, 2024.

[11] W. Cui, P. Zhao, J. Wang, N. Qin, E. A. Ho, and C. L. Ren, "Reagent free detection of sars-cov-2 using an antibody-based microwave sensor in a microfluidic platform," *Lab on a Chip*, vol. 22, no. 12, p. 2307–2314, 2022.

[12] M. R. Jamiruddin, B. A. Meghla, D. Z. Islam, T. A. Tisha, S. S. Khandker, M. U. Khondoker, M. A. Haq, N. Adnan, and M. Haque, "Microfluidics technology in sars-cov-2 diagnosis and beyond: A systematic review," *Life (Basel)*, vol. 12, no. 5, p. 649, 2022.

[13] H. Yousefi, A. Mahmud, D. Chang, J. Das, S. Gomis, J. B. Chen, H. Wang, T. Been, L. Yip, E. Coomes, *et al.*, "Detection of sars-cov-2 viral particles using direct, reagent-free electrochemical sensing," *Journal of the American Chemical Society*, vol. 143, no. 4, pp. 1722–1727, 2021.

[14] K. Ember, F. Daoust, M. Mahfoud, F. Dallaire, E. Z. Ahmad, T. Tran, A. Plante, M.-K. Diop, T. Nguyen, A. St-Georges-Robillard, *et al.*, "Saliva-based detection of covid-19 infection in a real-world setting using reagent-free raman spectroscopy and machine learning," *Journal of biomedical optics*, vol. 27, no. 2, pp. 025002–025002, 2022.

[15] R. Dorfman, "The detection of defective members of large populations," *Annals of Mathematical Statistics*, vol. 14, p. 436, 1943.

[16] L. Mutesa, P. Ndishimye, Y. Butera, J. Souopgui, A. Uwineza, R. Rutayisire, E. L. Ndoricimpaye, E. Musoni, N. Rujeni, T. Nyatanyi, *et al.*, "A pooled testing strategy for identifying sars-cov-2 at low prevalence," *Nature*, vol. 589, no. 7841, pp. 276–280, 2021.

[17] M. Aldridge and D. Ellis, "Pooled testing and its applications in the covid-19 pandemic," *Pandemics: Insurance and Social Protection*, pp. 217–249, 2022.

[18] J. Burtniak, A. Hedley, K. Dust, P. Van Caeseele, J. Bullard, and D. R. Stein, "Dorfman pooling enhances sars-cov-2 large-scale community testing efficiency," *PLOS Global Public Health*, vol. 3, no. 4, p. e0001793, 2023.

[19] T. N. Furstenau, J. H. Cocking, C. M. Hepp, and V. Y. Fofanov, "Sample pooling methods for efficient pathogen screening: Practical implications," *PLoS One*, vol. 15, no. 11, p. e0236849, 2020.

[20] J. Yi, R. Mudumbai, and W. Xu, "Low-cost and high-throughput testing of covid-19 viruses and antibodies via compressed sensing: System concepts and computational experiments," *arXiv preprint arXiv:2004.05759*, 2020.

[21] N. Shental, S. Levy, V. Wuvshet, S. Skorniakov, B. Shalem, A. Ottolenghi, Y. Greenshpan, R. Steinberg, A. Edri, R. Gillis, *et al.*, "Efficient high-throughput sars-cov-2 testing to detect asymptomatic carriers," *Science advances*, vol. 6, no. 37, p. eabc5961, 2020.

[22] S. Ghosh, A. Rajwade, S. Krishna, N. Gopalkrishnan, T. Schaus, A. Chakravarthy, S. Varahan, V. Appu, R. Ramakrishnan, S. Ch, M. Jindal, V. Bhupathi, A. Gupta, A. Jain, R. Agarwal, S. Pathak, M. Rehan, S. Consul, Y. Gupta, N. Gupta, P. Agarwal, R. Goyal, V. Sagar, U. Ramakrishnan, S. Krishna, P. Yin, D. Palakodeti, and M. Gopalkrishnan, "Tapestry: A single-round smart pooling technique for covid-19 testing," *medRxiv*, 2020.

[23] J. H. McDermott, D. Stoddard, P. J. Woolf, J. M. Ellingford, D. Gokhale, A. Taylor, L. A. Demain, W. G. Newman, and G. Black, "A nonadaptive combinatorial group testing strategy to facilitate health care worker screening during the severe acute respiratory syndrome coronavirus-2 (sars-cov-2) outbreak," *The Journal of Molecular Diagnostics*, vol. 23, no. 5, pp. 532–540, 2021.

[24] M. Escobar, G. Jeanneret, L. Bravo-Sánchez, A. Castillo, C. Gómez, D. Valderrama, M. Roa, J. Martínez, J. Madrid-Wolff, M. Cepeda, M. Guevara-Suarez, O. L. Sarmiento, A. Medaglia, M. Forero-Shelton, M. Velasco, J. M. Pedraza, R. Laajaj, S. Restrepo, and P. Arbelaez, "Smart pooling: Ai-powered covid-19 informative group testing," *Scientific Reports*, vol. 12, no. 1, p. 6519, 2022.

[25] J. Zhang and L. S. Heath, "Adaptive group testing strategy for infectious diseases using social contact graph partitions," *Scientific Reports*, vol. 13, no. 1, p. 12102, 2023.

[26] N. Barak, R. Ben-Ami, T. Sido, A. Perri, A. Shtoyer, M. Rivkin, T. Licht, A. Peretz, J. Magenheim, I. Fogel, *et al.*, "Lessons from applied large-scale pooling of 133,816 sars-cov-2 rt-pcr tests," *Science Translational Medicine*, vol. 13, no. 589, p. eabf2823, 2021.

[27] N. Lagopati, P. Tsioli, I. Mourkioti, A. Polyzou, A. Papaspyropoulos, A. Zafiropoulos, K. Evangelou, G. Sourvinos, and V. G. Gorgoulis, "Sample pooling strategies for sars-cov-2 detection," *Journal of virological methods*, vol. 289, p. 114044, 2021.

[28] A. Chakravarthy, S. Krishna, S. Ghosh, A. Tomar, S. Varahan, A. Rajwade, S. Ghosh, N. Gupta, R. Agarwal, H. Payal, P. Chakraborty, K. V. Vemula, A. Vyas, R. Goru, S. Krishna, D. Palakodeti, and M. Gopalkrishnan, "Large-scale testing for sars-cov-2 using tapestry pooling," *medRxiv*, 2020.

[29] I. A. Correa, T. de Souza Rodrigues, A. Queiroz, L. Nascimento, T. Wolff, R. Akamine, S. Kuriyama, L. da Costa, and A. Fidalogo-Neto, "Boosting sars-cov-2 detection combining pooling and multiplex strategies," *Scientific Reports*, vol. 12, p. 8684, 2022.

[30] K. S. Butler, B. D. Carson, J. D. Podlevsky, C. M. Mayes, J. M. Rowland, D. Campbell, J. B. Ricken, G. Wudiri, and J. A. Timlin, "Singleplex, multiplex and pooled sample real-time rt-pcr

assays for detection of sars-cov-2 in an occupational medicine setting," *Scientific Reports*, vol. 12, no. 1, p. 17733, 2022.

[31] N. L. Welch, M. Zhu, C. Hua, J. Weller, M. E. Mirhashemi, T. G. Nguyen, S. Mantena, M. R. Bauer, B. M. Shaw, C. M. Ackerman, *et al.*, "Multiplexed crispr-based microfluidic platform for clinical testing of respiratory viruses and identification of sars-cov-2 variants," *Nature medicine*, vol. 28, no. 5, pp. 1083–1094, 2022.

[32] D. Lee, C.-H. Chu, and A. F. Sarioglu, "Point-of-care toolkit for multiplex molecular diagnosis of sars-cov-2 and influenza a and b viruses," *Acs Sensors*, vol. 6, no. 9, pp. 3204–3213, 2021.

[33] K. J. Horadam, *Hadamard matrices and their applications*. Princeton university press, 2012.

[34] L. Ou, S. Liao, Z. Qin, and H. Yin, "Millimeter wave wireless hadamard image transmission for mimo enabled 5g and beyond," *IEEE Wireless Communications*, vol. 27, no. 6, pp. 134–139, 2020.

[35] A. Aung, B. P. Ng, and S. Rahardja, "Sequency-ordered complex hadamard transform: Properties, computational complexity and applications," *IEEE Transactions on Signal processing*, vol. 56, no. 8, pp. 3562–3571, 2008.

[36] D.-e. Jabeen, G. Monir, and F. Azim, "Sequency domain signal processing using complex hadamard transform," *Circuits, Systems, and Signal Processing*, vol. 35, pp. 1783–1793, 2016.

[37] C. Oliver and E. Pike, "Multiplex advantage in the detection of optical images in the photon noise limit," *Applied Optics*, vol. 13, no. 1, pp. 158–161, 1974.

[38] Z. Zhang, X. Wang, G. Zheng, and J. Zhong, "Hadamard single-pixel imaging versus fourier single-pixel imaging," *Optics Express*, vol. 25, no. 16, pp. 19619–19639, 2017.

[39] B. A. Yorke, G. S. Beddard, R. L. Owen, and A. R. Pearson, "Time-resolved crystallography using the hadamard transform," *Nature methods*, vol. 11, no. 11, pp. 1131–1134, 2014.

[40] P. Jacquinot, "The luminosity of spectrometers with prisms, gratings, or fabry-perot etalons," *JOSA*, vol. 44, no. 10, pp. 761–765, 1954.

[41] J. F. James and R. S. Sternberg, "The design of optical spectrometers," *London: Chapman and Hall*, 1969.

[42] G. S. Beddard and B. A. Yorke, "Pump–probe spectroscopy using the hadamard transform," *Applied Spectroscopy*, vol. 70, no. 8, pp. 1292–1299, 2016.

[43] Y. Zhang, M. O. Malik, J. Kang, C. Yuen, and Q. Liu, "Sequency encoding single pixel spectroscopy based on hadamard transform," *Optics Express*, vol. 30, no. 17, pp. 30121–30134, 2022.

[44] V. Yatsyna, A. H. Abikhodr, A. Ben Faleh, S. Warnke, and T. R. Rizzo, "High-throughput multiplexed infrared spectroscopy of ion mobility-separated species using hadamard transform," *Analytical Chemistry*, vol. 94, no. 6, pp. 2912–2917, 2022.

[45] Q. Xu, N. Ding, D. Ma, H. Lin, B. Lin, X. Ma, J. Yang, and L. Guo, "Portable hadamard-transform raman spectrometer: A powerful analytical tool for point-of-care testing," *Analytical Chemistry*, vol. 96, no. 30, pp. 12217–12224, 2024.

[46] M. Harwit and N. J. Sloane, *Hadamard Transform Optics*. Academic Press, 1979.

[47] H. Shani-Narkiss, O. D. Gilday, N. Yayon, and I. D. Landau, "Efficient and practical sample pooling for high-throughput pcr diagnosis of covid-19," *MedRxiv*, pp. 2020–04, 2020.

[48] A. Hedayat and W. D. Wallis, "Hadamard matrices and their applications," *The annals of statistics*, pp. 1184–1238, 1978.

1 Acknowledgements

We thank Ethan Beddard, Prof. Tia Keyes and Prof. Arwen Pearson for their constructive comments.

Data and code availability.

The code to calculated Hadmard matrices and perform the Hadamard pooling simulation is available on reasonable request.

Author Contributions

B.A.Y. data analysis, co-wrote the paper. G.S.B. research conception, code, data analysis, co-wrote the paper.

Competing Interests

The authors declare no competing interests.

Supporting Information

2 Appendix 1

Hadamard Matrices

A conventional Hadamard matrix is square and contains entries of 1 and -1 only, and has the property that each row is orthogonal and as such has matching entries in half the rows and half not. The first row and column contain only ones. A variation of these matrices is used here and are labelled as Hadamard **S** matrices in which the first row and column of the Hadamard matrix are ignored and the changes $1 \rightarrow 0$ and $-1 \rightarrow 1$ are made in the rest of the matrix. This is the form of the matrix shown in Figure 1.

Only certain integer values are allowed in forming the **S** matrices, nevertheless they are numerous, the first few are $3, 7, 11, 19, 23, 27, 31, 43, 47, 59, \cdots 103, \cdots, 199, \cdots$ which are prime numbers which also satisfy the condition $4n + 3$ where $n = 0, 1, 2, \cdots$. Other Hadamard matrices can be made with size $2n - 1, n = 2, 3$, for example $n = 4$ produces a 15×15 matrix and 27×27 by a similar method. The $n = 32$ and 31 S matrices are circulant. Their first rows are

$n = 23$, 0 0 0 0 1 0 1 0 0 1 1 0 0 1 1 0 1 0 1 1 1 1 1

$n = 31$, 0 0 1 0 0 1 0 0 0 0 1 1 1 0 1 0 1 0 0 0 1 1 1 1 0 1 1 0 1 1 1

The Hadamard Transform

If the situation arises that all samples have to be measured a vector is constructed out of every summed row in a Hadamard matrix, instead of just selected ones. This vector comprises the unknown values q_k which are multiplied by $S_i k = 0$ or 1 and summed according to the pattern of each row into:

$$W_i = \sum_k Q_k S_{ik} \tag{S1}$$

which is the dot product of vectors Q and a row of S. The column vector formed is $W = \begin{bmatrix} W_1, & W_2, & \cdots \end{bmatrix}^T$. The transform to return the original values is $Q = S^{-1}W$ where $Q = \begin{bmatrix} Q1, & Q2, & \cdots \end{bmatrix}$. The inverse of an S matrix can be calculated by directly, but more simply as

$$S^{-1} = \frac{2}{n+1}(2S^T - J) \tag{S2}$$

where J is an 'all ones' matrix and superscript T indicates the transpose.

$n = 7$

```
0 0 1 0 1 1 1
0 1 0 1 1 1 0
1 0 1 1 1 0 0
0 1 1 1 0 0 1
1 1 1 0 0 1 0
1 1 0 0 1 0 1
1 0 0 1 0 1 1
```

$n = 11$

```
0 1 0 0 0 1 1 1 0 1 1
1 0 0 0 1 1 1 0 1 1 0
0 0 0 1 1 1 0 1 1 0 1
0 0 1 1 1 0 1 1 0 1 0
0 1 1 1 0 1 1 0 1 0 0
1 1 1 0 1 1 0 1 0 0 0
1 1 0 1 1 0 1 0 0 0 1
1 0 1 1 0 1 0 0 0 1 1
0 1 1 0 1 0 0 0 1 1 1
1 1 0 1 0 0 0 1 1 1 0
1 0 1 0 0 0 1 1 1 0 1
```

$n = 15$

```
1 1 1 1 0 1 0 1 1 0 0 1 0 0 0
1 1 1 0 1 0 1 1 0 0 1 0 0 0 1
1 1 0 1 0 1 1 0 0 1 0 0 0 1 1
1 0 1 0 1 1 0 0 1 0 0 0 1 1 1
0 1 0 1 1 0 0 1 0 0 0 1 1 1 1
1 0 1 1 0 0 1 0 0 0 1 1 1 1 0
0 1 1 0 0 1 0 0 0 1 1 1 1 0 1
1 1 0 0 1 0 0 0 1 1 1 1 0 1 0
1 0 0 1 0 0 0 1 1 1 1 0 1 0 1
0 0 1 0 0 0 1 1 1 1 0 1 0 1 1
0 1 0 0 0 1 1 1 1 0 1 0 1 1 0
1 0 0 0 1 1 1 1 0 1 0 1 1 0 0
0 0 0 1 1 1 1 0 1 0 1 1 0 0 1
0 0 1 1 1 1 0 1 0 1 1 0 0 1 0
0 1 1 1 1 0 1 0 1 1 0 0 1 0 0
```

$n = 19$

```
0 1 1 0 0 0 0 1 0 1 0 1 1 1 1 0 0 1 1
1 1 0 0 0 0 1 0 1 0 1 1 1 1 0 0 1 1 0
1 0 0 0 0 1 0 1 0 1 1 1 1 0 0 1 1 0 1
0 0 0 0 1 0 1 0 1 1 1 1 0 0 1 1 0 1 1
0 0 0 1 0 1 0 1 1 1 1 0 0 1 1 0 1 1 0
0 0 1 0 1 0 1 1 1 1 0 0 1 1 0 1 1 0 0
0 1 0 1 0 1 1 1 1 0 0 1 1 0 1 1 0 0 0
1 0 1 0 1 1 1 1 0 0 1 1 0 1 1 0 0 0 0
0 1 0 1 1 1 1 0 0 1 1 0 1 1 0 0 0 0 1
1 0 1 1 1 1 0 0 1 1 0 1 1 0 0 0 0 1 0
0 1 1 1 1 0 0 1 1 0 1 1 0 0 0 0 1 0 1
1 1 1 1 0 0 1 1 0 1 1 0 0 0 0 1 0 1 0
1 1 1 0 0 1 1 0 1 1 0 0 0 0 1 0 1 0 1
1 1 0 0 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1
1 0 0 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 1
0 0 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 1 1
0 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 1 1 0
1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 1 1 0 0
1 0 1 1 0 0 0 0 1 0 1 0 1 1 1 1 0 0 1
```

$n = 27$

```
0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0
1 1 0 1 1 0 0 1 1 1 1 0 0 0 0 0 0 0 0 1 1 1 1 0 0 1 1
0 0 0 1 0 0 1 1 0 1 0 0 1 0 1 0 1 1 0 1 0 0 1 1 0 1 0
1 1 1 1 0 1 1 0 0 1 1 1 1 0 0 0 0 0 0 0 0 1 1 1 1 0 0
0 1 0 0 0 1 0 0 1 1 0 1 0 0 1 0 1 0 1 0 1 1 0 1 0 0 1
1 0 0 1 1 1 0 1 1 0 0 1 1 1 1 0 0 0 0 0 0 0 0 1 1 1 1
0 0 1 1 0 0 0 1 0 0 1 1 0 1 0 0 1 0 1 0 1 0 1 1 0 1 0
1 1 1 0 0 1 1 1 0 1 1 0 0 1 1 1 1 0 0 0 0 0 0 0 0 1 1
0 1 0 0 1 1 0 0 0 1 0 0 1 1 0 1 0 0 1 0 1 0 1 0 1 1 0
1 1 1 1 1 0 0 1 1 1 0 1 1 0 0 1 1 1 1 0 0 0 0 0 0 0 0
0 1 0 1 0 0 1 1 0 0 0 1 0 0 1 1 0 1 0 0 1 0 1 0 1 0 1
1 0 0 1 1 1 1 0 0 1 1 1 0 1 1 0 0 1 1 1 1 0 0 0 0 0 0
0 0 1 0 1 1 0 1 0 0 1 1 0 0 0 1 0 0 1 1 0 1 0 0 1 0 1
1 0 0 0 0 1 1 1 1 0 0 1 1 1 0 1 1 0 0 1 1 1 1 0 0 0 0
0 0 1 0 1 1 0 1 0 0 1 1 0 0 0 1 0 0 1 1 0 1 0 0 1 0 1
1 0 0 0 0 0 0 0 0 1 1 1 1 0 0 1 1 1 0 1 1 0 0 1 1 1 1
0 0 1 0 1 0 1 0 1 1 0 1 0 0 1 1 0 0 0 1 0 0 1 1 0 1 0
1 1 1 0 0 0 0 0 0 0 0 1 1 1 1 0 0 1 1 1 0 1 1 0 0 1 1
0 1 0 0 1 0 1 0 1 0 1 1 0 1 0 0 1 1 0 0 0 1 0 0 1 1 0
1 1 1 1 0 0 0 0 0 0 0 0 1 1 1 1 0 0 1 1 0 1 1 0 0 1 1
0 1 0 1 0 0 1 0 1 0 1 1 0 1 0 0 1 1 0 0 0 1 0 0 1 0 1
1 0 0 1 1 1 1 0 0 0 0 0 0 0 0 1 1 1 1 0 0 1 1 1 0 1 1
0 0 1 1 0 1 0 0 1 0 1 0 1 0 1 1 0 1 0 0 1 1 0 0 0 1 0
1 1 1 0 0 1 1 1 1 0 0 0 0 0 0 0 0 1 1 1 1 0 0 1 1 1 0
0 1 0 0 1 1 0 1 0 0 1 0 1 0 1 0 1 1 0 1 0 0 1 1 0 0 0
```

Fig. 5: Figure S-1. Examples of Hadamard **S** matrices. Each of the matrices except $n = 27$ is circulant and so can be produced from the first lines alone by rotating by one position. Note that the 15×15 matrix cannot be made by the Quadratic Residue method, however, different but functionally equivalent matrices are formed by either the Shift Register method (matrix shown above) or the Doubling method, both of these methods are described in [46] and the 28×28 Hadamard matrix construction, among others [33, 48]