

Contact Tracing Enhances the Efficiency of COVID-19 Group Testing

Ritesh Goenka,^{*i} Shu-Jie Cao,^{*s}
Chau-Wai Wong,ⁿ Ajit Rajwade,ⁱ Dror Baronⁿ

ⁱ IIT Bombay, India

^s ShanghaiTech University, China

ⁿ North Carolina State University, USA



NC STATE
UNIVERSITY

CISS 2021

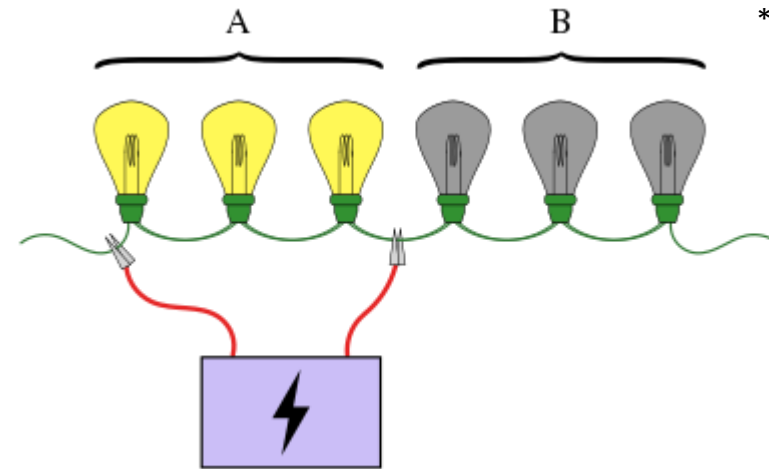
* Equal Contributions.

Conventional Testing for COVID-19

- Conventional testing steps:
 - Collect sample using nasal or oropharyngeal swab.
 - Amplify genetic material with reverse transcription polymerase chain reaction (RT-PCR).
 - Report positive/negative based on thresholding amplified genetic material.
- Challenges:
 - Resource intense.
 - False negatives & false positives.
- Want more efficient testing.

Group Testing to Reduce # of Tests Needed

- Goal: Test a population of size n with fewer tests $m (< n)$.
- Initial idea [Dorfman'43]^Δ:
 - Test individuals in groups of a given size, e.g., 3.
 - A group tested negative → all healthy.
 - Tested **positive** → continue with individual testing.
- Limitations of Dorfman's approach:
 - Assumes i.i.d. health status.
 - Fragile to false negatives & positives.



^Δ Dorfman, "The detection of defective members of large populations," 1943.

* Figure reproduced from https://en.wikipedia.org/wiki/Group_testing

Recent Progress of Group Testing Applied to COVID-19

- Can optimize group size [1].
- Compressed sensing formulation with prevalence rate, solved via message-passing style algorithm [2].
- Modeling multiplicative noise, and end-to-end study of decoding in pooled qPCR – including matrix design, prevalence rate estimation [3].
- Exploited individual's symptom and family structure [4].
- Exploited community structure with a focus on encoder design [5–6], but did not consider noise.

[1] Hanel and Thurner, “Boosting test-efficiency by pooled testing strategies for SARS-CoV-2,” Mar. 2020.

[2] Zhu, Rivera, and Baron, “Noisy pooled PCR for virus testing,” Apr. 2020.

[3] Ghosh et al., “A compressed sensing approach to group testing for COVID-19 detection,” May 2020.

[4] Zhu, Rivera, Rush, and Baron, “Noisy pooled PCR for COVID-19 testing,” May 2020.

[5] Nikolopoulos, Guo, Fragouli, and Diggavi, “Community aware group testing,” Jul. 2020.

[6] Nikolopoulos, Srinivasavaradhan, Guo, Fragouli, and Diggavi, “Group testing for overlapping communities,” Dec. 2020.

Group Testing: Compressed Sensing Approach

$$\begin{matrix} m \times 1 \\ \mathbf{y} \end{matrix} = \text{noise} \left(\begin{matrix} m \times n, \underline{m} < n \\ \mathbf{A} \end{matrix} \begin{matrix} n \times 1 \\ \mathbf{x} \end{matrix} \right)$$

Vector of **observed quantities** giving health status of m groups, $m < n$.

Known sensing matrix:
 $A_{ij} = 1$ if j^{th} sample contributes to i^{th} group, else 0.

Sparse vector of **unknown** health status of n individuals.

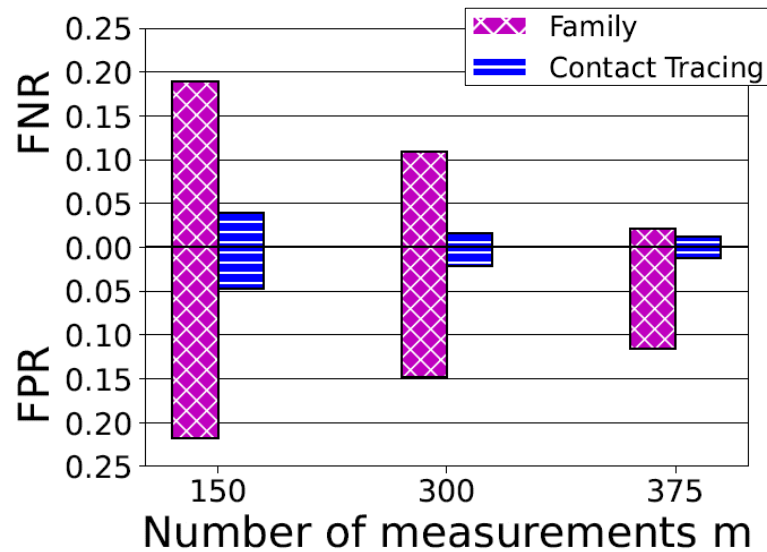
- Estimate \mathbf{x} given \mathbf{y} , \mathbf{A} , noise model.
- Use of different decoding algorithms (message passing,* LASSO#)
- Single-stage (nonadaptive), unlike Dorfman.
- Can incorporate side information into the probability distribution of \mathbf{x} .

* Zhu et al., "Noisy pooled PCR for virus testing," Apr. 2020.

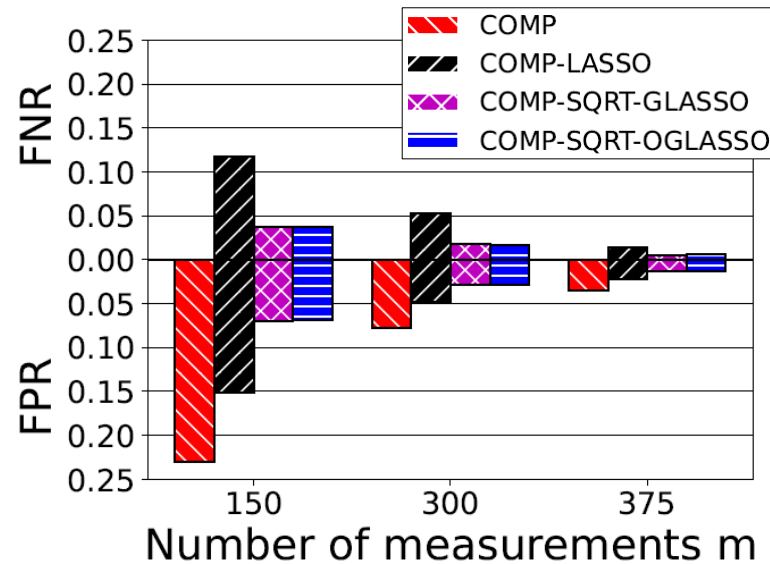
Ghosh et al., "A compressed sensing approach to group testing for COVID-19 detection," May 2020.

Our Contributions

- Improved test efficiency by using *side information (SI)*:
 - Family structure: nonoverlapping or overlapping.
 - Contact-tracing data: (i) who's in contact, (ii) physical proximity, and (iii) contact duration.



Proposed method for binary noise

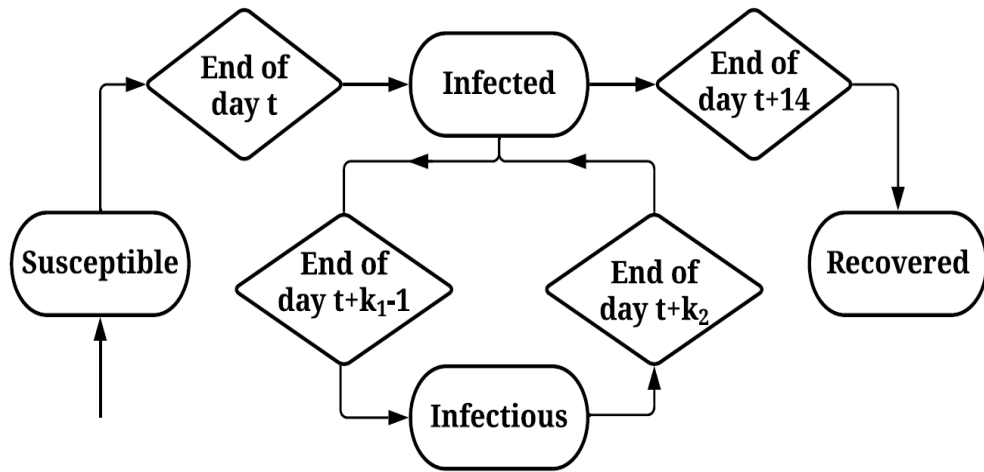


Proposed method for multiplicative noise

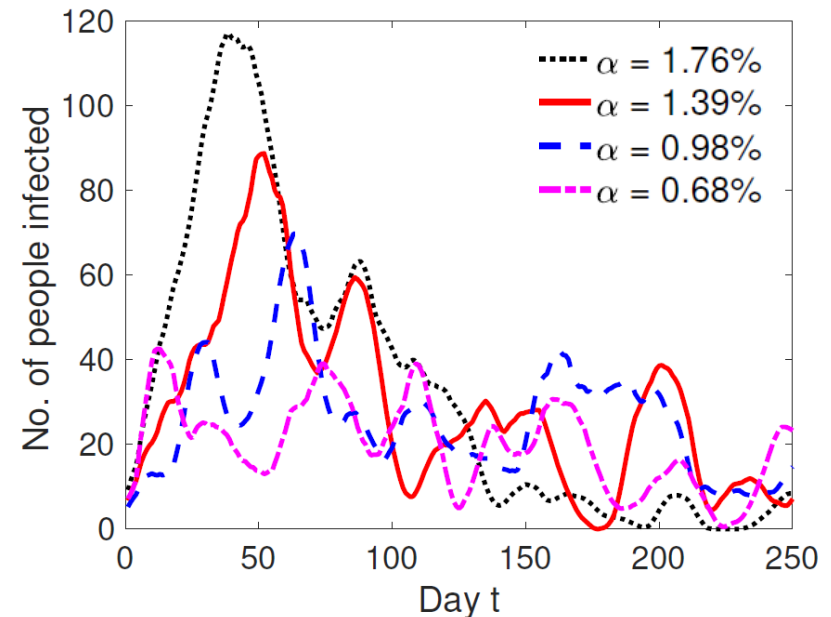
4% sparsity + only
15% measurements:
Both methods can
achieve
~ 5% FNR & FPR.

Simulate Infected Population w/ Contract-Tracing Info

- Use SEIR* style **generative infection model** incorporating contact-tracing side information (SI) for data simulation at individual level.



State transition diagram for an individual



Number of active infections

* Carcione et al., "A simulation of a COVID-19 epidemic based on a deterministic SEIR model", 2020.

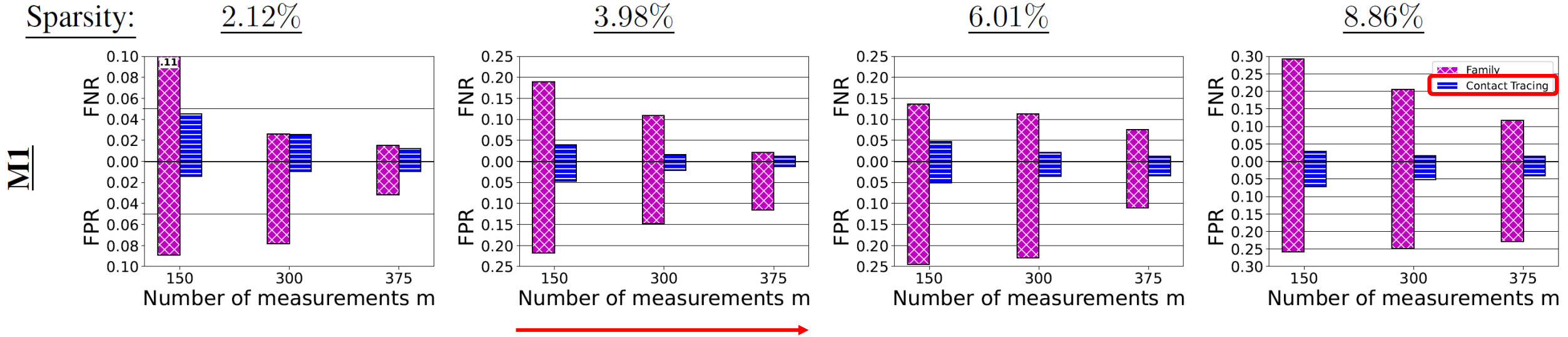
Binary Noise (M1): Proposed Method

- Binary noise model (M1):
 - Pooling: $\mathbf{w} = \mathbf{A}\mathbf{x}$, $\mathbf{A} \in \{0,1\}^{m \times n}$, $\mathbf{x} \in \{0,1\}^n$.
 - Erroneous PCR testing: $\mathbb{P}(y_i = 1 | w_i = 0) = 0.1\%$, $\mathbb{P}(y_i = 0 | w_i \neq 0) = 2\%$.
- Decoding: Generalized approximate message passing (GAMP)* framework.
- Key: The design of **denoiser**, $g_{\text{in}}(\mathbf{v}) = \mathbb{E}[X_i | \mathbf{V} = \mathbf{v}]$, where \mathbf{v} is pseudo data.
 - Family denoiser #: Individual's group membership information.
 - Contract-tracing denoiser: (i) Who. (ii) How far. (iii) How long.

J. Zhu, K. Rivera, C. Rush, and D. Baron, "Noisy pooled PCR for COVID-19 testing," *Paris Machine Learning Meetup*, May 2020.

* S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," *IEEE Int. Symp. Inf. Theory*, 2011.

Binary Noise (M1): Numerical Results



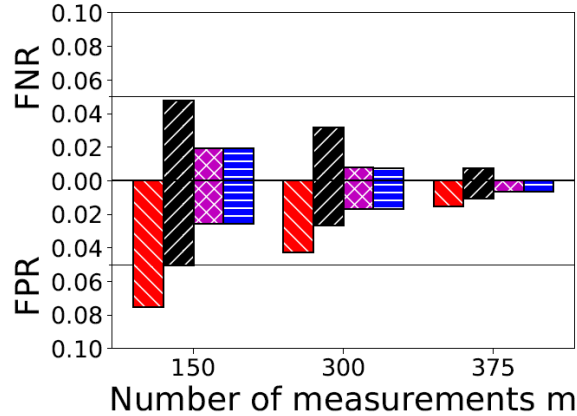
- Contact-tracing data as SI helps more than family structure.
- The larger the measurement rate, m/n , the better the performance.
- Using contract-tracing data, FNR & FNR < 5% except for challenges cases of sparsity level = 8.86%.

Multiplicative Noise (M2): Proposed Method

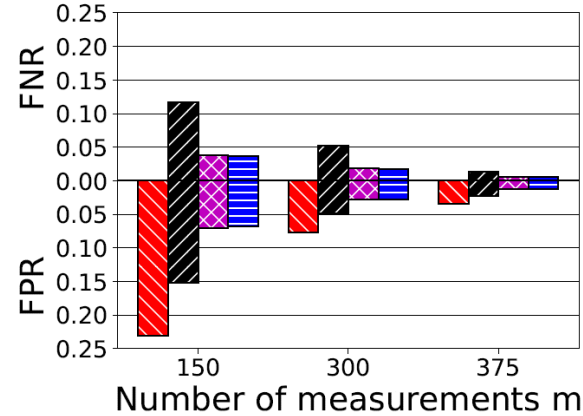
- Multiplicative noise model (M2):
 - Pooling: $\mathbf{w} = \mathbf{A}\mathbf{x}$, $\mathbf{A} \in \{0,1\}^{m \times n}$, $\mathbf{x} \in [0, \infty)^n$.
 - Noisy RT-PCR amplification: $\mathbf{y} = \mathbf{w} \circ (1 + q)^{\mathcal{N}(\mathbf{0}, \sigma^2 \mathbb{I})}$.
 $q \in (0,1]$: known amplification factor, $\sigma^2 \ll 1$ controls strength of PCR noise.
- Decoding: Group Lasso based algorithms.
 - Family structure: group square-root Lasso.
$$\hat{\mathbf{x}}^{\text{SQRT-GLASSO}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 + \rho \sum_{g=1}^{n_1} \|\mathbf{x}_g\|_2$$
 - Contact-tracing data: overlapping group square-root Lasso. Preprocessed by clique detection. Cost function in similar form.

Multiplicative Noise (M2): Numerical Results

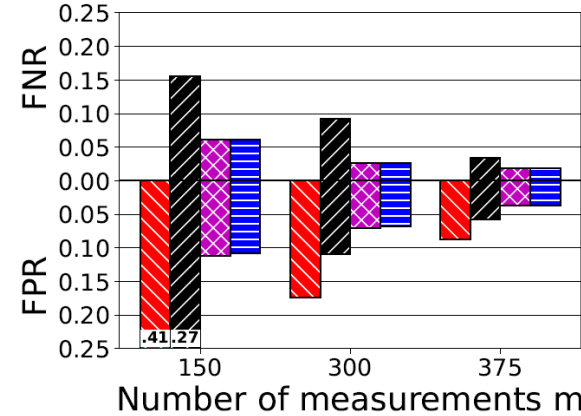
Sparsity: 2.12%



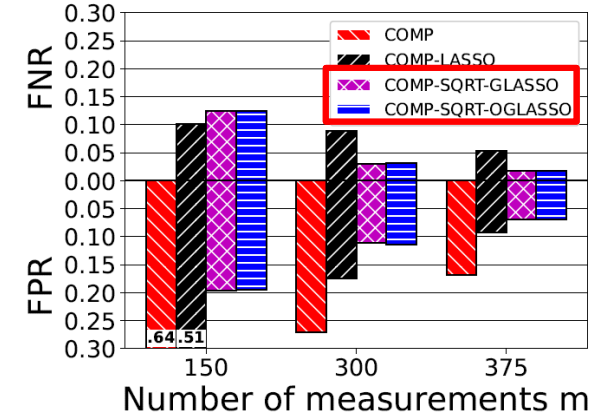
Sparsity: 3.98%



Sparsity: 6.01%



Sparsity: 8.86%



- Using family/contact-tracing SI significantly improves the performance over decoding without SI.
- M2 can allow estimating viral loads of infected individuals.
- Both models are robust to inaccurate specification of contact-tracing information (duration/proximity of contact info. from Bluetooth).

Conclusion and Future Directions

- Improved efficiency by exploiting ***family structure*** and ***contact tracing***.
- Have achieved ~5% FNR & FPR at 4% sparsity level/prevalence rate with the need of only 15% of tests required in a conventional testing scenario.
- Future Directions:
 - Design better group testing matrices by leveraging the insights from coding community.
 - Calibrate an exact number of groups for required performance.
 - Explicitly model RT-PCR noise.