$x^{(l)} \sim f_{q,\mathcal{G},\epsilon}$, i.i.d., let $\widehat{\pi}^{\epsilon}(\mathcal{G})$ be the *empirical proportion of ϵ-MSNE*: $\widehat{\pi}^{\epsilon}(\mathcal{G}) \equiv \frac{1}{m}\sum_{l=1}^{m} \mathbb{1}\left[x^{(l)} \in \mathcal{NE}^{\epsilon}(\mathcal{G})\right]$. We denote the *Kullback-Leibler (KL) divergence* between two Bernoulli distributions with parameters $p_1, p_2 \in (0,1)$ by $\mathrm{KL}(p_1 \| p_2)$.

Proposition 2.2. **(Maximum-likelihood Estimation)** *The tuple* $(\hat{\mathcal{G}}, \widehat{q}, \widehat{\epsilon})$ *is a* maximum likelihood estimator (MLE), *with respect to dataset D, for the parameters of the generative model* $f_{(q,\mathcal{G},\epsilon)}$, *as defined in Eqn. 1 if and only if (iff)* $\widehat{q} = \widehat{\pi}^{\widehat{\epsilon}}(\widehat{\mathcal{G}})$, *and* $(\widehat{\mathcal{G}}, \widehat{\epsilon}) \in \arg\max_{(\mathcal{G},\epsilon)} \mathrm{KL}(\widehat{\pi}^{\epsilon}(\mathcal{G}) \| \pi^{\epsilon}(\mathcal{G}))$.

Dealing with $\pi^{\epsilon}(\mathcal{G})$ directly would require us to compute all ϵ-MSNE of $\mathcal{G}$; computing only one ϵ-MSNE is PPAD-hard in general [3, 4]. The following lemma provides bounds on the KL divergence.

Lemma 2.3. *Given a game* $\mathcal{G}$ *with* $0 < \pi^{\epsilon}(\mathcal{G}) < \widehat{\pi}^{\epsilon}(\mathcal{G})$ *and* $\mu(\mathcal{NE}^{\epsilon}(\mathcal{G})) \in (0,1)$, *we have*

$$-\widehat{\pi}^{\epsilon}(\mathcal{G})\log\pi^{\epsilon}(\mathcal{G}) - \log 2 < \mathrm{KL}(\widehat{\pi}^{\epsilon}(\mathcal{G}) \| \pi^{\epsilon}(\mathcal{G})) < -\widehat{\pi}^{\epsilon}(\mathcal{G})\log\pi^{\epsilon}(\mathcal{G}) .$$

From the above, it is easy to see that when $\pi^{\epsilon}(\mathcal{G})$ is "low enough," we can obtain an approximation to the MLE by simply maximizing $\widehat{\pi}^{\epsilon}(\mathcal{G})$ only: i.e., $\arg\max_{\mathcal{G}} \mathrm{KL}(\widehat{\pi}^{\epsilon}(\mathcal{G}) \| \pi^{\epsilon}(\mathcal{G})) \approx \arg\max_{\mathcal{G}} \widehat{\pi}^{\epsilon}(\mathcal{G})$.

3 APPLICATION: GENERALIZED IDS GAMES

In α-IDS games [2] with n state-agents, each state-agent i determines whether or not to invest in protection (against epidemics). We denote $a_i = 1$ if i invests and $a_i = 0$ if i does not invest and let x_i be the probability that $a_i = 1$. We let $x = (x_1, ..., x_n)$ to be the joint-mixed strategy profile of all agents and x_{-S} to be the profile of all agents that are not in S. There is a cost of investment C_i and loss L_i associated with the bad event occurring, either through a direct or indirect (transferred) contamination. We denote by p_i the probability that agent i will experience the bad event from a direct contamination and by q_{ji} the probability that agent i will experience the bad event due to transfer exposure from agent j. The parameter $\alpha_i \in [0,1]$ specifies the probability that agent i's investment will not protect i against transfers of a bad event. Given the parameters, the *expected cost function of agent i* is $M_i(x_i, x_{-i})$

$$\equiv x_i[C_i + \alpha_i r_i(x_{-i})L_i] + (1 - x_i)[p_i + (1 - p_i)r_i(x_{-i})]L_i ,$$

where $r_i(x_{-i}) \equiv 1 - s_i(x_{-i})$ and $s_i(x_{-i}) \equiv \prod_{j \neq i}(x_j + (1 - x_j)(1 - q_{ji}))$ are i's overall risk and safety functions, respectively. By definition, an ϵ-MSNE x of an α-IDS game satisfies

$$M_i(x_i, x_{-i}) - \epsilon \le M_i(0, x_{-i}) \;\&\; M_i(x_i, x_{-i}) - \epsilon \le M_i(1, x_{-i}). \quad (2)$$

Learning. We approximate our MLE objective by maximizing the number of ϵ-MSNE in the data when the true proportion of ϵ-MSNE of the game is less than the empirical proportion of ϵ-MSNE in the dataset *We empirically observe that the true proportion of ϵ-MSNE in α-IDS games is very low.* This would justify Lemma 2.3 and our method.

We subdivide the optimization by first optimizing over $\mathcal{G}$, and then optimizing over ϵ. We use an upper bound by applying Eqn. 2. Then, we approximate the indicator function in the upper bound using a sigmoid function, which is the standard approach leading to the BackProp algorithm in neural networks [11].

Using standard primal-dual optimization and regularization techniques, we obtain and solve a non-linear program (using gradient-ascent/descent optimization) subject to the respective constraints on the variables. The process terminates when the objective function satisfies some condition and after exceeding some threshold based on the total running time (i.e., $\approx$ 5 hours for the CDC dataset).

4 EXPERIMENTS ON VACCINATION DATA

Viewing each State as a player in the game, we interpret the vaccination percentages as mixed-strategies and generate 1500 samples i.i.d. according to an n-variate jointly-independent Gaussian PDF, where $n = 48$, with the joint mean and standard deviations given by each State's reported vaccination rate and standard deviation in the CDC 2009-2010 US States H1N1 data [1]. This is our way to account for the noise in the data. We impose an *a priori* bias for learning where only neighboring states may transfer the virus.

Learned α-IDS Games. Although game parameters themselves are not our main interest, we highlight similar observations on 10 learned games because they provide anecdotal validation.

Players' Characteristics. All of the players have strategic substitutability behavior – this happens if $\alpha_i < 1 - p_i$ for each player i. In Figure 1, the x-axis denotes the α values of the players, the y-axis denotes the $1 - p$ values of the players, and the line is the equation $\alpha = 1 - p$. The plot is scaled to capture the α and $1 - p$ values. The plot illustrates that our learning formulation produces values of the parameters that are consistent with vaccination scenarios, in which $\alpha < 1 - p$.

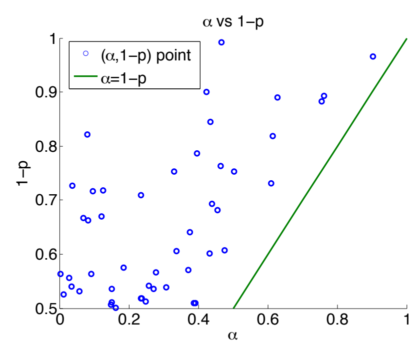

Figure 1: Players' Characteristics of a Learned Game.

Player's Best-Response Correspondences. All of the players have non-trivial best-response – players do not have any "obvious" dominant strategies.

Players' Transfer Risks. The transfer risks of the players are not random – they are correlated to the training examples.

Players' Equilibrium Behavior. Given the learned games, we run a version of some learning-heuristics/regret-minimization [8], in which we use the mean vaccination rates as the initial mixed-strategy profile to compute ϵ-MSNE in these games.

It turns out that the mean vaccination-rates given in the CDC data is an 0.35-MSNE of the learned game. We are able to find an exact MSNE which is also a PSNE after trying many initial mixed-strategies that are drawn uniformly at random.

[1]https://www.cdc.gov/flu/fluvaxview/reportshtml/reporti0910/resources/2009-10_coverage.xlsx

REFERENCES

[1] Robert G. Bartle. 1995. *The Elements of Integration and Lebesgue Measure.* Wiley-Interscience.

[2] H. Chan and L. E. Ortiz. 2014. Computing Nash Equilibria in Generalized Interdependent Security Games. In *Advances in Neural Information Processing Systems 27 (NIPS'14).* 2735–2743.

[3] X. Chen, X. Deng, and S. Teng. 2009. Settling the Complexity of Computing Two-player Nash Equilibria. *J. ACM* 56, 3, Article 14 (May 2009), 57 pages.

[4] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. 2006. The Complexity of Computing a Nash Equilibrium. In *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing (STOC '06).* 71–78.

[5] Q. Duong, Y. Vorobeychik, S. Singh, and M. P. Wellman. 2009. Learning Graphical Game Models. In *Proceedings of the 21st International Jont Conference on Artifical Intelligence (IJCAI'09).* San Francisco, CA, USA, 116–121.

[6] Q. Duong, M. P. Wellman, S. Singh, and M. Kearns. 2012. Learning and Predicting Dynamic Networked Behavior with Graphical Multiagent Models. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 1 (AAMAS '12).* 441–448.

[7] S. Ficici, D. Parkes, and A. Pfeffer. 2008. Learning and Solving Many-Player Games through a Cluster-Based Representation. In *Proceedings of the Twenty-Fourth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI'08).* 188–195.

[8] D. Fudenberg and D. K. Levine. 1998. *The Theory of Learning in Games.* MIT Press Books, Vol. 1. MIT Press.

[9] D. Fudenberg and J. Tirole. 1991. *Game Theory.* MIT Press, Cambridge, MA.

[10] X. A. Gao and A. Pfeffer. 2010. Learning Game Representations from Data Using Rationality Constraints. In *Proceedings of the 26th Conference on on Uncertainty in Artificial Intelligence (UAI'10).*

[11] S. Haykin. 1999. *Neural Networks: A Comprehensive Foundation.* Prentice Hall.

[12] J. Honorio and L. E. Ortiz. 2015. Learning the Structure and Parameters of Large-Population Graphical Games from Behavioral Data. *Journal of Machine Learning Research.* 16, 1 (June 2015), 1157–1210.

[13] M. Kearns and J. Wortman. 2008. Learning from Collective Behavior. In *In Proceedings of the 21st Annual Conference on Learning Theory (COLT '04).* 99–110.

[14] J. Nash. 1951. Non-Cooperative Games. *Annals of Mathematics* 54 (September 1951), 286–295. Issue 2.

[15] Y. Vorobeychik, M. P. Wellman, and S. Singh. 2007. Learning Payoff Functions in Infinite Games. *Mach. Learn.* 67, 1-2 (May 2007), 145–168.

[16] J. Wright and K. Leyton-Brown. 2010. Beyond Equilibrium: Predicting Human Behavior in Normal-Form Games. In *Proceedings of the National Conference on Artificial Intelligence (AAAI '10).* 901–907.

[17] J. Wright and K. Leyton-Brown. 2012. Behavioral Game Theoretic Models: A Bayesian Framework for Parameter Analysis. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 2 (AAMAS '12).* 921–930.

[18] B. D. Ziebart, J. A. Bagnell, and A. K. Dey. 2010. Modeling Interaction via the Principle of Maximum Causal Entropy. In *Proceedings of the 27th International Conference on Machine Learning (ICML '10).* 1255–1262.