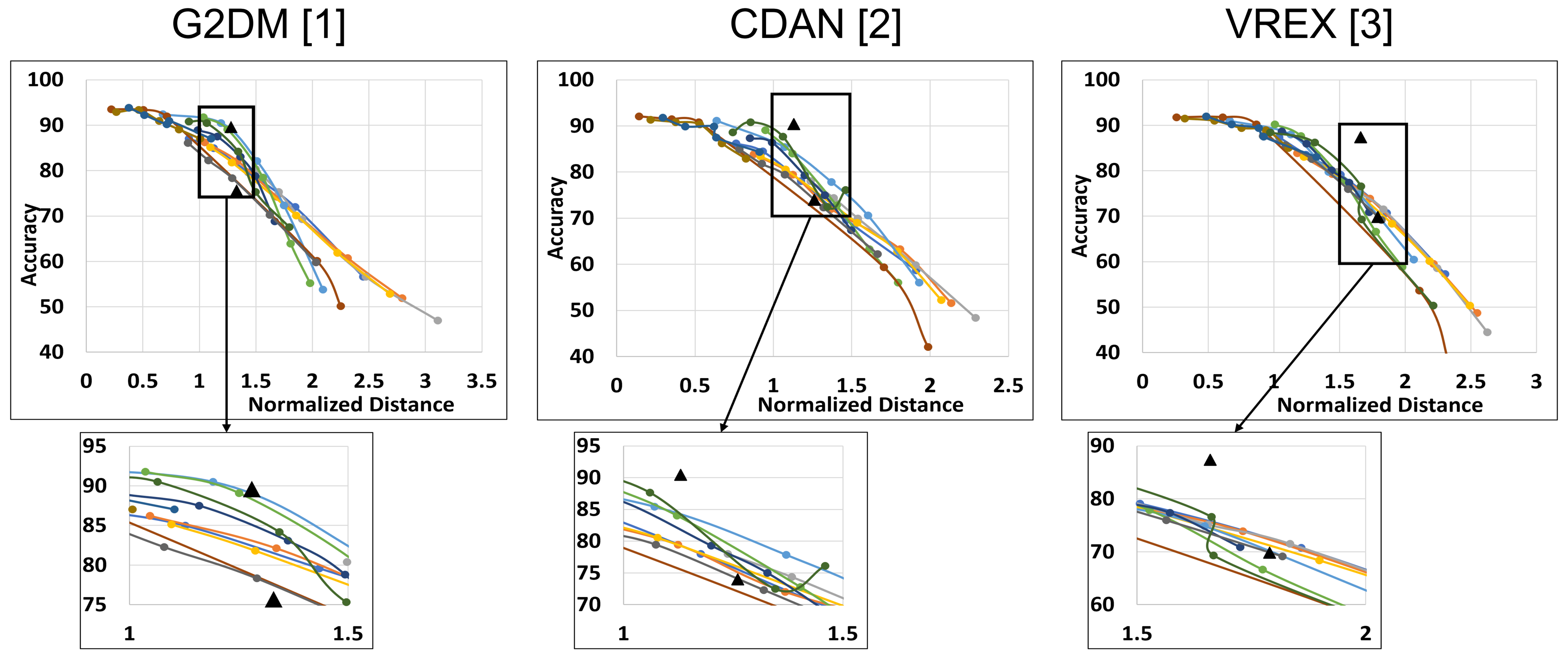
Do Domain Generalization Methods Generalize Well?

Akshay Mehra[1], Bhavya Kailkhura[2], Pin-Yu Chen[3] and Jihun Hamm[1]

[1]Tulane University, [2]Lawrence Livermore National Laboratory, [3]IBM Research

Motivation

- Domain Generalization (DG) focuses on developing models that generalize well to data from domains unseen during training.
- Following works that explain the generalization performance using distributional distance, learning a representation space that reduces distance between domains is shown to be effective at DG.
- However, evaluating DG models only using benchmark datasets is insufficient to gauge their performance on unseen domains.



Contributions

- We propose a data independent and a distance-based evaluation method for DG based on distributionally robust optimization.
- Our method *efficiently* estimates the loss of the worst-case distribution to better gauge the generalization performance of DG models. It can be easily incorporated into training of DG models to produce models that generalize better on unseen domains.

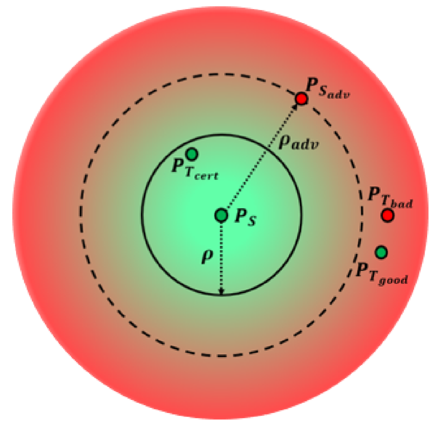
Distance-based evaluation method for DG

- Notation: $\mathcal{X}$, $\mathcal{Y}$ denotes the data domain and the labels, $g: \mathcal{X} \to \mathcal{Z}$ denotes the representation map and $h: \mathcal{Z} \to \mathcal{Y}$ denotes the classifier on top of $\mathcal{Z}$.
- Our distance-based measure of the performance of DG model relies on the worst-case loss of the model at a particular distance $\sup_{\mathcal{P}: \mathcal{W}_2(\mathcal{P},\mathcal{Q}) \leq \rho} \mathbb{E}_{(x,y)\sim\mathcal{P}}[\ell(h(g(x)),y)]$. This is an infinite-dimensional problem over a convex set and is difficult to solve.
- However, its optimal value can be computed using the dual problem. Moreover, since DG methods learn a representation space where unseen domains lie close to the source domain, we solve the following to compute the worst-case loss

$$\sup_{\mathcal{P}: \mathcal{W}_2(\mathcal{P}, g\#\mathcal{Q}) \leq \rho} \mathbb{E}_{(z,y)\sim\mathcal{P}}[\ell(h(z),y)]$$

$$= \inf_{\gamma \geq 0} \left\{ \gamma\rho^2 + \mathbb{E}_{(z_0,y_0)\sim g\#Q}\left[\sup_{z\in\mathcal{Z}}\{\ell(h(z),y) - \gamma \|z_0 - z\|_2^2\}\right]\right\}.$$

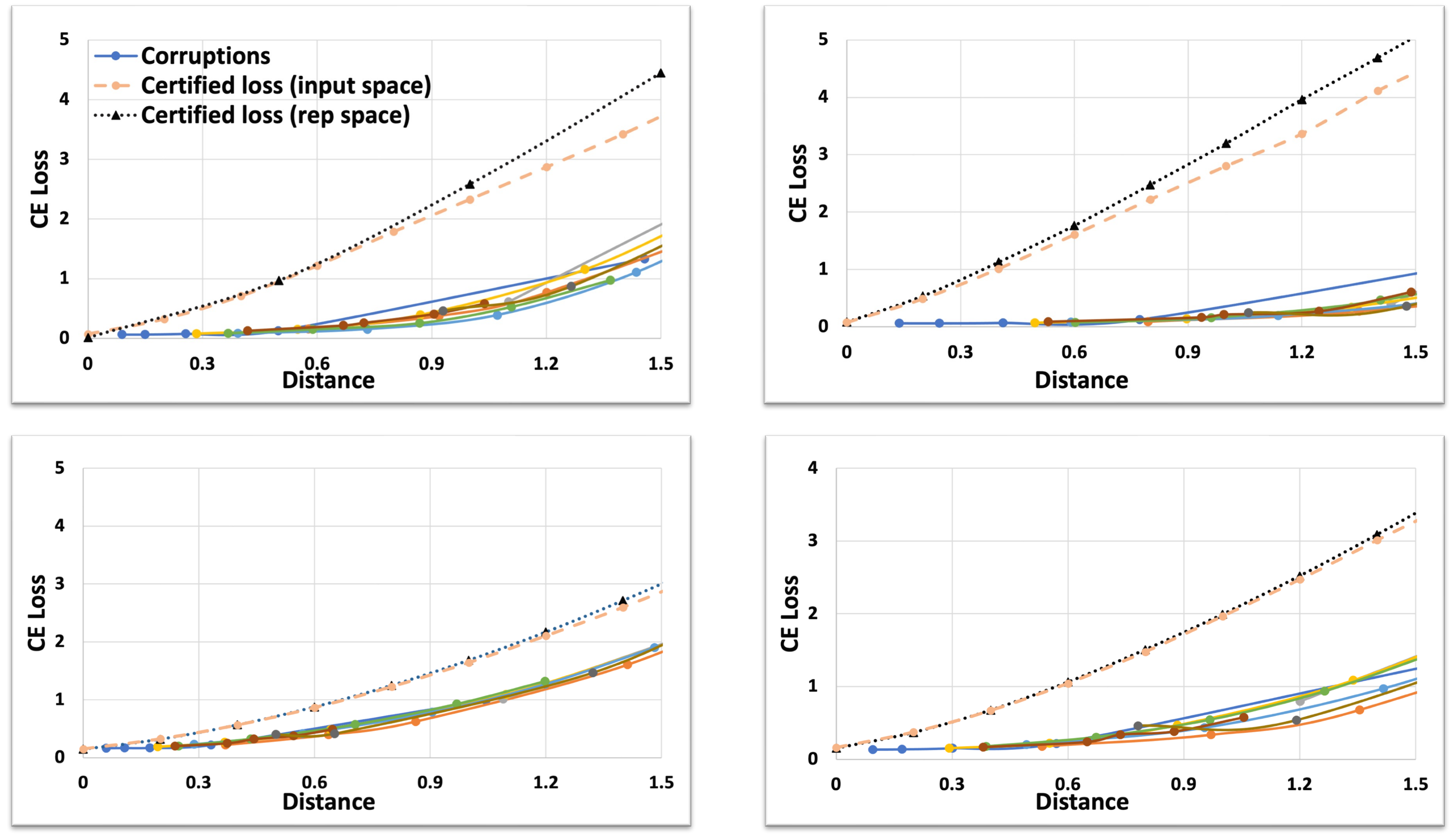
- The space of probability distributions with Wasserstein distance as a metric. $\mathcal{P}_S$ and $\mathcal{P}_T$ are the sources and the unseen target distributions. The worst-case loss can be efficiently computed for any distance ρ by solving the above problem.



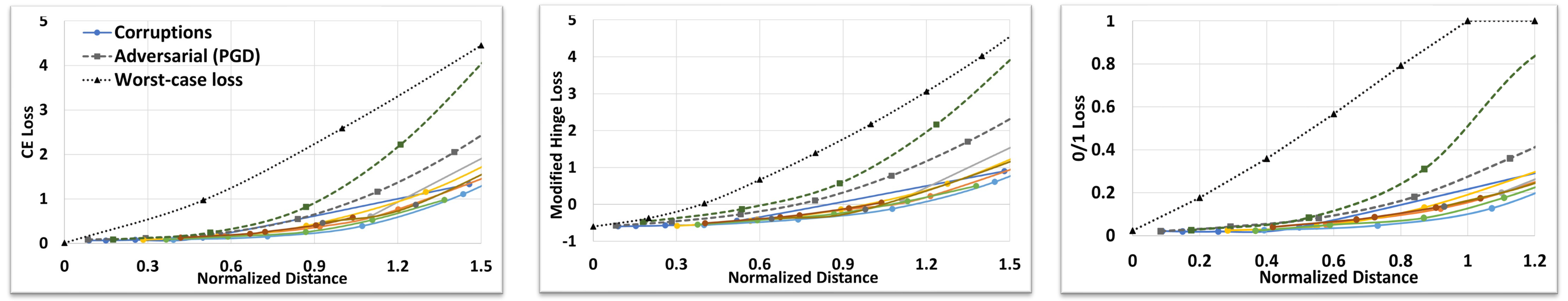
- The objective of minimizing the worst-case loss of a DG model at any distance in the representation space can be combined with losses of other DG methods to yield models that generalize better on unseen domains.

Key results

- High worst-case loss even close to the source domains highlights the poor generalizability of current DG models.
- Worst-case loss in the representation space is not an overestimation of the worst-case loss in the input space.
  - Models trained with WM & G2DM on R-MNIST.



- Effectiveness of using different loss functions (cross entropy, hinge and misclassification) for computing the worst-case loss.
  - Models trained with WM on R-MNIST dataset and are evaluated using WC-DG.



References

- Albuquerque et. al (2019), Generalizing to unseen domains via distribution matching.
- Long et. al (2018), Conditional adversarial domain adaptation.
- Krueger et. al (2021), Out-of-distribution generalization via risk extrapolation (rex).
- Sinha et. al (2017) Certifying some distributional robustness with principled adversarial training.
- Volpi et al (2018) Generalizing to unseen domains via adversarial data augmentation.

This work was supported by the NSF EPSCoR-Louisiana Materials Design Alliance (LAMDA) program #OIA-1946231 and by the LLNL-LDRD Program under Project No. 20-ERD-014. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.