Adversarial Robustness of Streaming Algorithms through Importance Sampling

**Model**
- **Input**: Elements of an underlying data set $S$, which arrives sequentially and adversarially. Adversary can choose future inputs after seeing previous outputs by honest algorithm.
- **Output**: Evaluation (or approximation) of a given function.
- **Goal**: Use space sublinear in the size $m$ of the input $S$.

Surprising separation between “classic” streaming model where the stream input is fixed but the order of the updates may be given adversarially.

Hardt and Woodruff [HW13] showed that linear sketches are NOT robust to adversarial attacks, must use $\Omega(n)$ space by giving an attack on AMS $F_2$ algorithm.

**Applications / Motivations**
- Adversarial machine learning: ML problems where the input is chosen by an adversary.
- Database queries: For multiple queries to a database, each query may depend on the responses to the previous queries.
- Transparency of Algorithms: Internal state of honest algorithms may be entirely revealed or otherwise compromised.

**Coresets**
- **Coreset**: Returns an $\epsilon$-approximation on a query space.
- **Merge and reduce framework**: Each $C_i$ is an $\epsilon \log n$-coreset of the corresponding partition of the substream.

**Applications**: $k$-means clustering, $k$-median clustering, projective clustering, principal component analysis, Bayesian logistic regression, generative adversarial networks, $k$-line center, $M$-estimators.

**Corollary**: Merge-and-reduce is adversarially robust.

**Intuition**: Importance is a robust metric and adversarial attacks must change it.

Vladimir Braverman (Johns Hopkins University, Google), Avinatan Hassidim (Google), Yossi Matias (Google), Mariano Schain (Google), Sandeep Silwal (MIT), Samson Zhou (Carnegie Mellon University)

---

**Empirical Evaluations**
- **Streaming $k$-means clustering**: a series of point batches where all points except the last batch are randomly sampled from a two-dimensional standard normal distribution. Points in the last batch sampled but around a distant center.
- **Streaming linear regression**: all batches except the last one are sampled around a constellation of four points in the plane such that the optimal regression line is of $-1$ slope through the origin. The last batch is at $(L, L)$, far from the origin so the resulting optimal regression line has slope $1$ through the origin.
- **Sampling vs. sketching**: For a random unit sketching matrix $S$ (each of its elements is sampled from $\{-1,1\}$ with equal probability), we create an adversarial data stream $M$ such that its columns are in the nullspace of $S$ for linear regression.

---

**References**