Strategic Noise – Why Different

- **Stochastic Noise** (errors ~ Distribution)
  - any error, s.t., loss < gain
- **Strategic Noise** (any error, s.t., loss < gain)
  - Can report: \((x_i, y_i) \in \mathbb{R}^{d+1}\)
  - Single Peaked Preferences: prefer outcome \(y_i\)

**Model**

- \(n\) agents, each controls one datapoint: \((x_i, y_i) \in \mathbb{R}^{d+1}\)
- \(x_i\): public information, \(y_i\): private (manipulable)
- Agents are strategic:
  - Can report: \((x_i, y_i) \in \mathbb{R}^{d+1}\) s.t., \(y_i \neq y_i\)
  - Single Peaked Preferences: prefer outcome \(y_i\)

**Our Goal.** Construct Linear Regression mechanism \(M^*(Y)\):

- Learn relationship between \(x, y\)
- Induce strategyproofness (SP) (prefer truth telling from misreporting, irrespective of reports of others) without monetary incentives

### Example of a SP Linear Regression (corrected) CRM Family [Perote & Perote-Peña, 2004]

1) Split dataset into \(S, S'\) s.t.: \(S = S'\) or \(S', S'\': \) separable
2) From each point in \(S\), compute CWA for points in \(S'\)
3) Final line: median-of-median CWAs
   - Only defined for 2D, not generalizable to higher dimensions

### Generalized Resistant Lines (GRL)

#### Definition \([S, S', k, k'] - GRL\). Choose \(S, S'\) separable and \(k \in [I(S)], k' \in [I(S')]\). Output line \((\beta_k, \beta_0)\) s.t.:
\[
\min_{y \in S} (y - \beta_k x - \beta_0) = \min_{y' \in S'} (y' - \beta_k x - \beta_0) = 0
\]

E.g., previous example was a \((S, S', 2, 2)\)-GRL mechanism

**Generalized Resistant Hyperplanes (GRH)**

- Choose \(S_1, \ldots, S_{d+1}\) publicly separable and \(k_1 \in [I(S_1)], \ldots, k_{d+1} \in [I(S_{d+1})]\). Output line \((\beta_{d+1}, \ldots, \beta_0)\):
\[
\min_{y \in S_1} (y - \beta_{d+1} x - \beta_0) = \min_{y' \in S_{d+1}} (y' - \beta_{d+1} x - \beta_0) = 0
\]

#### Theorem. GRH mechanisms yield unique solution and are group-strategyproof.

**Proof Idea.**
- Any coalition creates a new hyperplane.
- Uniqueness of GRH for given \(k_1, \ldots, k_{d+1}\) new hyperplane either does not exist, or some agent is not rational

### Efficiency of SP Linear Regressors

When no strategic considerations: Ordinary Least Squares (OLS) is most popular, but

**In fact, CRM \(\subseteq\) GRL**

GRL mechanisms generalize to higher dimensions.

**Problems:**
1) Definition of separability in higher dimensions
2) Uniqueness of solution

**Ham Sandwich Theorem**

**Theorem** (Stone & Tukey, 1942). Given \(k\) continuous measures \(\mu_1, \ldots, \mu_k\) on \(\mathbb{R}^d\), \(\exists\) hyperplane, \(H: \mu_i(H^+) = \frac{1}{2}, \forall i \in [k]\).

- **Discrete Variant** (still bisecting) due to [Elton & Hill, 2011]
- **Unique Variant** due to [Steiger and Zhao, 2010]
  \(\rightarrow\) close to what we need

### Open Questions

1) What about consistency?
2) SP linear regressors for other types of agent incentives?
3) Constructive characterization of all SP linear regressors.

**Overarching Goal**

To build a Theory of Incentives for ML algorithms.