

## A Dual Information-Geometric Perspective

- loss minimization focuses on **function** computed by AdaBoost (i.e., weights on **weak classifiers**)
- **dual view**: instead focus on **distributions**  $D_t$  (i.e., weights on **examples**)
- dual perspective combines **geometry** and **information theory**
- exposes underlying mathematical **structure**
- basis for proving **convergence**