Modulator Design for Binary Classification of Poisson Measurements

Jiaji Huang, Robert Calderbank
Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA

Introduction

What we want to do:

Poisson channel $\mathbf{y} \sim$ pois$(\mathbf{Ax}), \mathbf{y} \in \mathbb{Z}_+^m, \mathbf{A} \in \mathbb{R}_+^{m\times n}, \mathbf{x} \in \mathbb{R}_+^n$

Two hypotheses $H_1 : \mathbf{x} = \mathbf{u}$

$H_2 : \mathbf{x} = \mathbf{v}$

Classification Given $\mathbf{y}$, decide $\mathbf{x}$ as $\mathbf{u}$ or $\mathbf{v}$ without recovering it

Goal design an improved system yielding smaller classification error

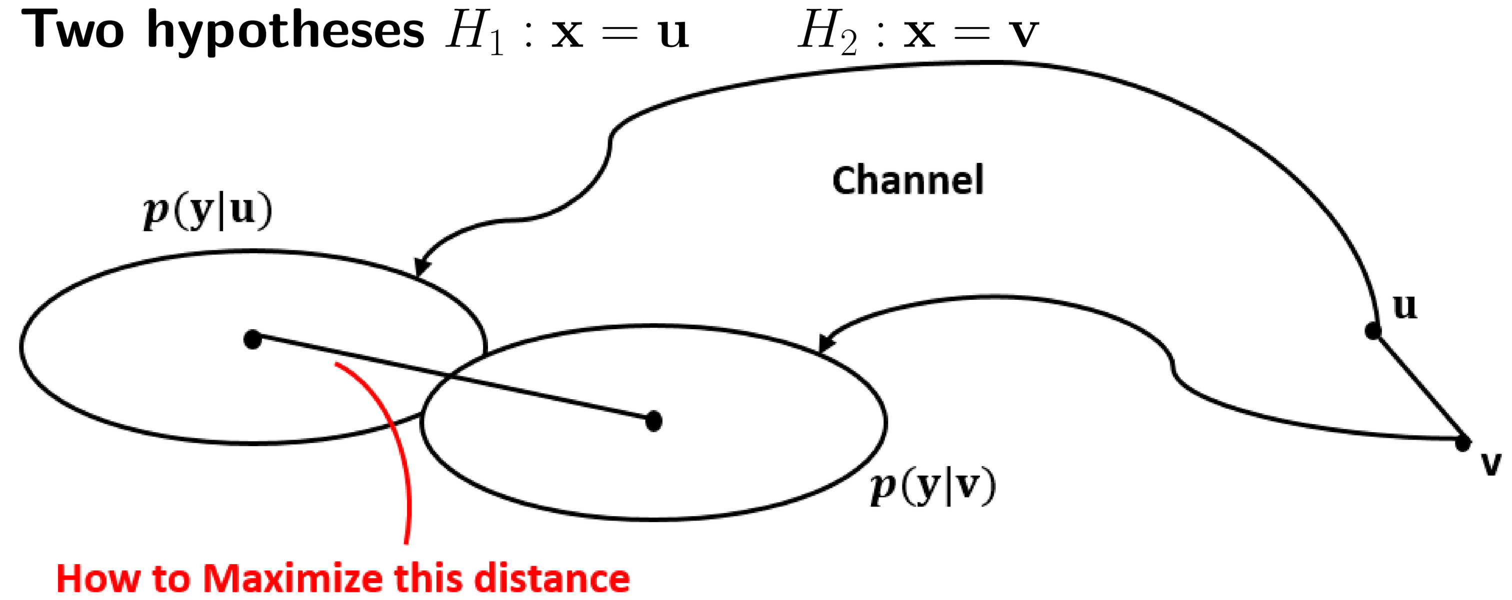
Keyword: Binary Classification, Bayesian Classifier, Diversity, Classification Error

Motivation

Consider a Gaussian Channel:

Measurements $\mathbf{y} = \mathbf{Ax} + \mathbf{n}$

Two hypotheses $H_1 : \mathbf{x} = \mathbf{u}$ $\quad H_2 : \mathbf{x} = \mathbf{v}$



Solution Rotate $\mathbf{u} - \mathbf{v}$ to the dominant eigenvector of $\mathbf{A}^\top\mathbf{A}$

Back to Poisson case:

Idea Design a modulator $\mathbf{R}$ that "rotates" $\mathbf{u} - \mathbf{v}$ appropriately

designed system $\mathbf{y} \sim$ pois$(\mathbf{ARx})$

Problem Setup

Error of optimal classifier

Hypothesis testing:

$$H_1 : \mathbf{y} \sim \text{pois}(\mathbf{ARu})$$
$$H_2 : \mathbf{y} \sim \text{pois}(\mathbf{ARv})$$

Bayesian(optimal) classifier:

$$\frac{\text{pois}(\mathbf{y}|\mathbf{ARu})}{\text{pois}(\mathbf{y}|\mathbf{ARv})} \underset{H_2}{\overset{H_1}{\gtrless}} 1$$

Classification error:

$$\begin{aligned} P_e &= \sum_{\mathbf{y}} \min\{\text{pois}(\mathbf{y}|\mathbf{ARu}), \text{pois}(\mathbf{y}|\mathbf{ARv})\} \\ &\le \exp\{-\frac{1}{2}\sum_i \left(\sqrt{(\mathbf{ARu})_i} - \sqrt{(\mathbf{ARv})_i}\right)^2\} \\ &\triangleq P_e^{UB} \end{aligned}$$

Formulation

$$\begin{aligned} &\max_{\mathbf{R}} \sum_i \left(\sqrt{(\mathbf{ARu})_i} - \sqrt{(\mathbf{ARv})_i}\right)^2 \\ &\text{s.t.} \quad 0 \le R_{i,j} \le 1 \end{aligned} \qquad \text{(P0)}$$

A relaxed problem

Difficulty of problem P0

Convex maximization: optima occurs at vertex

Enumerating: infeasible, complexity $O(2^{n^2})$

A relaxed problem

Assmue $(\mathbf{ARu})_i + (\mathbf{ARv})_i \le C, \ \forall i = 1, \ldots, m$

Relaxed error upper bound

$$\tilde{P}_e^{UB} = \exp\left\{-\frac{1}{4C}\|\mathbf{ARu} - \mathbf{ARv}\|^2\right\}$$

Relaxed problem

$$\begin{aligned} &\max_{\mathbf{R}} \|\mathbf{ARu} - \mathbf{ARv}\|^2 \\ &\text{s.t.} \quad 0 \le R_{i,j} \le 1 \end{aligned} \qquad \text{(P1)}$$

Proposition Denote $\mathbf{r} = \text{vec}(\mathbf{R}^\top)$, $\boldsymbol{\gamma} = \text{vec}(\mathbf{R})$, $\mathbf{d} = \mathbf{u} - \mathbf{v}$

$$\|\mathbf{ARu} - \mathbf{ARv}\|^2 = \mathbf{r}^\top \underbrace{\left[(\mathbf{A}^\top\mathbf{A}) \otimes (\mathbf{dd}^\top)\right]}_{\mathbf{D}} \mathbf{r} = \boldsymbol{\gamma}^\top \underbrace{\left[(\mathbf{dd}^\top) \otimes (\mathbf{A}^\top\mathbf{A})\right]}_{\mathbf{E}} \boldsymbol{\gamma}$$

P1 as a fixed rank 0-1 quadratic maximization problem

Search the vertices of a zonotope, complexity $O(n^{2rank(\mathbf{D})})$

P1 as a mincut problem

Let $\boldsymbol{\rho} = 2\mathbf{r} - \mathbf{1}$, then $\rho_i = \pm 1$

$$\mathbf{r}^\top\mathbf{D}\mathbf{r} = \frac{1}{4}\underbrace{[1\ \boldsymbol{\rho}^\top]}_{\mathbf{s}^\top}\underbrace{\begin{bmatrix}\mathbf{1}^\top\mathbf{D}\mathbf{1} & \mathbf{1}^\top\mathbf{D} \\ \mathbf{D}\mathbf{1} & \mathbf{D}\end{bmatrix}}_{\mathbf{G}}\begin{bmatrix}1 \\ \boldsymbol{\rho}\end{bmatrix} = -\frac{1}{2}\sum_{i,j:s_i=-s_j} G_{i,j} + \text{const.}$$

Nontrivial if there exists any negative edge, $G_{i,j} < 0$

Even all edges are nonnegative, complexity is at least $O(n^4 \log^3 n)$

Fast Algorithm

Decompose the objective w.r.t. $\mathbf{R}$'s rows:

$$\mathbf{r}^\top\mathbf{D}\mathbf{r} = \mathbf{r}_1^\top[\|\mathbf{a}_1\|^2\mathbf{dd}^\top]\mathbf{r}_1 + (\mathbf{r}_2^\top[\|\mathbf{a}_2\|^2\mathbf{dd}^\top]\mathbf{r}_2 + 2[(\mathbf{a}_1^\top\mathbf{a}_2)\mathbf{r}_1^\top\mathbf{dd}^\top]\mathbf{r}_2) + \ldots$$

Algorithm 1 Fast Algorithm for problem P1

Input: Two hypothetical inputs: $\mathbf{u}, \mathbf{v}$. Sensing matrix $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$
Output: binary matrix $\mathbf{R} = [\mathbf{r}_1, \ldots, \mathbf{r}_n]^\top$

```
d = u − v
if 1ᵀd > 0 then
  r₁ ← ½[sign(d) + 1]
else
  r₁ ← ½[sign(−d) + 1]
end if
for k = 2, ..., n do
  if (‖a_k‖² 1ᵀ + 2 Σ_{i=1}^{k−1} (a_iᵀ a_k) r_iᵀ) d > 0 then
    r_k ← ½[sign(d) + 1]
  else
    r_k ← ½[sign(−d) + 1]
  end if
end for
```

Complexity $O(n^3)$

Extension: Adding Flux-preserving Constraint

$$\begin{aligned} &\max_{\mathbf{R}} \|\mathbf{ARu} - \mathbf{ARv}\|^2 \\ &\text{s.t.} \quad 0 \le R_{i,j} \le 1, \sum_i R_{i,j} \le 1 \end{aligned} \qquad \text{(P2)}$$

Decompose the objective w.r.t. $\mathbf{R}$'s columns:

$$\boldsymbol{\gamma}^\top\mathbf{E}\boldsymbol{\gamma} = d_1^2\boldsymbol{\gamma}_1^\top\mathbf{A}^\top\mathbf{A}\boldsymbol{\gamma}_1 + (d_2^2\boldsymbol{\gamma}_2^\top\mathbf{A}^\top\mathbf{A}\boldsymbol{\gamma}_2 + 2(d_1 d_2)\boldsymbol{\gamma}_1^\top(\mathbf{A}^\top\mathbf{A})\boldsymbol{\gamma}_2) + \ldots$$

Algorithm 2 Fast Algorithm for problem P2

Input: Two hypothetical inputs: $\mathbf{u}, \mathbf{v}$. Sensing matrix $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$
Output: binary matrix $\mathbf{R} = [\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_n]$

$\mathbf{d} = \mathbf{u} - \mathbf{v}$
$\boldsymbol{\gamma}_{1,i} = 1$ where $i = \arg\max_i \|\mathbf{a}_i\|^2$
for i=2,...,n **do**
$\boldsymbol{\gamma}_i = \mathbf{e}_k$ where $\mathbf{e}_k$ is the k-th natural basis such that

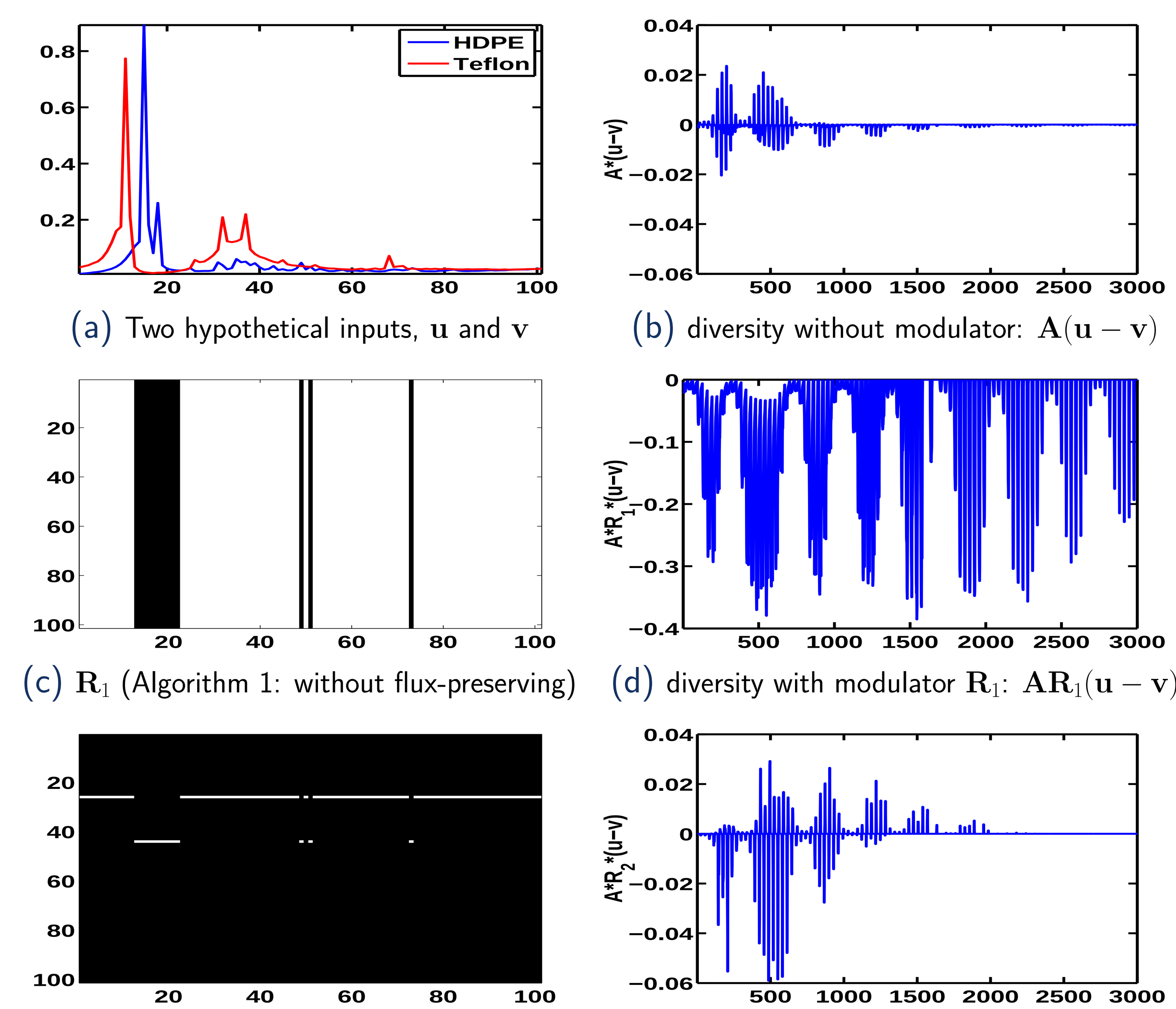
$$k = \arg\max_k d_i^2\mathbf{e}_k^\top\mathbf{A}^\top\mathbf{A}\mathbf{e}_k + 2\left[\sum_{j=1}^{i-1}(d_j d_i)\boldsymbol{\gamma}_j^\top(\mathbf{A}^\top\mathbf{A})\right]\mathbf{e}_k$$

end for

Complexity $O(n^3)$

Example

Real data $\mathbf{u}, \mathbf{v} \in \mathbb{R}_+^{101}$, $\mathbf{A} \in \mathbb{R}_+^{3040\times 101}$



(a) Two hypothetical inputs, $\mathbf{u}$ and $\mathbf{v}$
(b) diversity without modulator: $\mathbf{A}(\mathbf{u}-\mathbf{v})$
(c) $\mathbf{R}_1$ (Algorithm 1: without flux-preserving)
(d) diversity with modulator $\mathbf{R}_1$: $\mathbf{AR}_1(\mathbf{u}-\mathbf{v})$
(e) $\mathbf{R}_2$ (Algorithm 2: with flux-preserving)
(f) diversity with modulator $\mathbf{R}_2$: $\mathbf{AR}_2(\mathbf{u}-\mathbf{v})$

Figure: Hypotheses(a), modulator(c,e) and diversity of measurements(b,d,f)

Table: Classification Accuracy

no modulator	modulator $\mathbf{R}_1$	modulator $\mathbf{R}_2$
76.8%	100.0%	82.6%