

# Hybrid Algorithms for Generating Optimal Discrimination Designs

Ping-Yang Chen<sup>1</sup> Ray-Bing Chen<sup>1</sup> Cheng-Lin Hsu<sup>1</sup> Weng Kee Wong<sup>2</sup>

Department of Statistics, National Cheng Kung University, Tainan, Taiwan

Department of Biostatistics, University of California, Los Angeles, Los Angeles, CA, USA

## SUMMARY

Finding T- or KL-optimal discrimination designs is generally a difficult task because these design criteria contain multiple layers of nested optimization. We propose hybrid algorithms based on particle swarm optimization (PSO) to solve the complicated optimization problems. When there are only two competing models to discriminate, we propose a hybrid algorithm that combines PSO and a Newton type algorithm for the search. When there are three or more competing models, we propose a multi-layer algorithm to optimize the max-min T- and KL-optimality criteria. As an application, we apply our techniques to discriminate among five nonlinear dose-response models in toxicology. We also compare the performances of our algorithms with currently existing algorithms.

## BACKGROUND

- Consider a class of  $K$  possible models,  $f_1(y | x, \theta_1, \sigma^2), \dots, f_K(y | x, \theta_K, \sigma^2)$ 
  - $f_k$  is the  $k$ th probability distribution function for the response variable  $y$ .
  - $x$  is the independent variable from a design space  $\mathcal{X}$ .
  - $\theta_k \in \Theta_k$  is an unknown parameter vector of the  $k$ th model.
  - $\sigma^2$  is a nuisance parameter (e.g. the variance of the probability distribution).

- An approximate design is defined by

$$\xi = \{s_1, s_2, \dots, s_n; p_1, p_2, \dots, p_n\}$$

where  $s_j \in \mathcal{X}$  is the  $j$ th support points with corresponding weight  $p_j$  satisfying  $\sum_j p_j = 1$ .

- The goal is to find a design  $\xi^*$  that maximizes some discrimination criterion,  $\Phi(\xi)$ .

## PARTICLE SWARM OPTIMIZATION (PSO)

To find the optimal design, PSO is used because it is

- Assumption-free to work on non-differentiable optimal criterion  $\Phi(\xi)$ .
- Easy to develop by two updating steps with linear combinations involved only.

PSO first generates a swarm of particles,  $\xi_1^{(0)}, \dots, \xi_N^{(0)}$  and then updates them by two key steps:

$$V_i^{(t)} = \omega^{(t)} V_i^{(t-1)} + c_1 R_1 \otimes [\xi_{l(i)}^{(t-1)} - \xi_i^{(t-1)}] + c_2 R_2 \otimes [\xi_g^{(t-1)} - \xi_i^{(t-1)}],$$

$$\xi_i^{(t)} = \xi_i^{(t-1)} + V_i^{(t)}, \text{ for } i = 1, \dots, N$$

where

- $\xi_{l(i)}^{(t-1)}$  is the currently best position found by the  $i$ th particle.
- $\xi_g^{(t-1)}$  is the currently global best position found by all particles.
- $R_1$  and  $R_2$  are random vectors, and,  $c_1$ ,  $c_2$  and  $\omega^{(t)}$  are tuning parameters.

For maximin optimization problems, such as the optimal discrimination design problems, we need to modify PSO for more optimization layers.

## CRITERION FOR $K = 2$ AND PSO-QN ALGORITHM

- WLOG, suppose the true model  $f_t = f_1$  with known parameter  $\theta_t$ .
- The rival model is  $f_2$  whose parameter  $\theta_2 \in \Theta_2$ .
- The measure is the Kullback-Leibler (KL) divergence between  $f_t$  and  $f_2$ ,

$$J(f_t, f_2, x, \theta_2) = \int f_t(y | x, \sigma^2) \log \left\{ \frac{f_t(y | x, \sigma^2)}{f_2(y | x, \theta_2, \sigma^2)} \right\} dy, x \in \mathcal{X}.$$

E.g. For homoscedastic and normal errors, i.e. T-optimality in Atkinson and Fedorov (1975),  $J(f_t, f_2, x, \theta_2) = [f_t(x) - f_2(x, \theta_2)]^2$ .

- López-Fidalgo et al. (2007): The KL-optimal criterion is

$$\max_{\xi \in \Xi} \Phi(\xi) = \max_{\xi \in \Xi} I_{2,t}(\xi) = \max_{\xi \in \Xi} \min_{\theta_2 \in \Theta_2} \left\{ \int_{\mathcal{X}} J(f_t, f_2, x, \theta_2) \xi(dx) \right\}. \quad (1)$$

- The KL-optimal design  $\xi^*$  maximizes (1).
- Our PSO-QN algorithm has **two** optimization layers:
  - The non-differentiable Outer Loop is handled by PSO.
  - The differentiable Inner Loop is handled by the Newton-type L-BFGS algorithm.

## CRITERION FOR $K > 2$ AND PSO-S-QN ALGORITHM

- First identify KL-optimal designs  $\xi_k^*$  for each competing pair,  $f_t$  and  $f_k$ ,  $2 \leq k \leq K$ , and obtain their criterion values  $I_{2,t}(\xi_2^*), \dots, I_{K,t}(\xi_K^*)$ .
- Tommasi et al. (2016): The max-min KL-optimal criterion is

$$\max_{\xi \in \Xi} \Phi(\xi) = \max_{\xi \in \Xi} I_m(\xi) = \max_{\xi \in \Xi} \min_{2 \leq k \leq K} \frac{I_{k,t}(\xi)}{I_{k,t}(\xi_k^*)}. \quad (2)$$

- The max-min KL-optimal design  $\xi_m^*$  maximizes (2).
- Our PSO-S-QN algorithm has **three** optimization layers:
  - The non-differentiable Outer Loop is handled by PSO.
  - The Middle Loop is a screening procedure for  $K - 1$  efficiency values.
  - The differentiable Inner Loops for all  $I_{k,t}(\xi)$ ,  $2 \leq k \leq K$ , are handled by L-BFGS algorithm.

## TOXICOLOGY EXAMPLE

- Five toxicology models defined on  $x \in [0, 1250]$ ,

$$\begin{aligned} v_1(x, \theta_1) &= a, & v_2(x, \theta_2) &= a e^{-x/b}, \\ v_3(x, \theta_3) &= a e^{-(x/b)^d}, & v_4(x, \theta_4) &= a \left( c - (c - 1) e^{-x/b} \right), \\ v_5(x, \theta_5) &= a \left( c - (c - 1) e^{-(x/b)^d} \right), \end{aligned}$$

- Set  $v_5$  to be the true model with known  $(a, b, c, d) = (4.28, 835.57, 0.74, 3.52)$ .

- Assume errors are log-normally distributed.

First, PSO-QN finds the KL-optimal designs for each competing pair,  $v_5$  vs.  $v_k$ ,  $k = 1, 2, 3, 4$ .

Table: Comparison among algorithms for the case  $v_5$  vs.  $v_4$ . Summaries for 50 replications.

Search Algorithm	KL-Efficiency				CPU Time (seconds)
	Min.	Max.	#(90%+) <sup>†</sup>	#(100%) <sup>‡</sup>	
PSO-QN	<b>99.55</b>	100.00	<b>50</b>	<b>49</b>	<b>90.42</b>
NestedPSO <sup>1</sup>	14.49	100.00	45	35	231.85
Fedorov-Wynn <sup>2</sup>	0.00	96.78	48	0	114.72

<sup>1</sup>Chen et al. (2015); <sup>2</sup>Atkinson and Fedorov (1975)

<sup>†</sup>the number of resulting designs with at least 90% KL-efficiency.

<sup>‡</sup>the number of resulting KL-optimal designs.

Next, PSO-S-QN finds the max-min KL-optimal design, that is,

$$\xi_m = \{0.000, 451.530, 1043.591, 1250.000; 0.223, 0.342, 0.248, 0.188\}$$

The equivalence theorem (Tommasi et al., 2016)

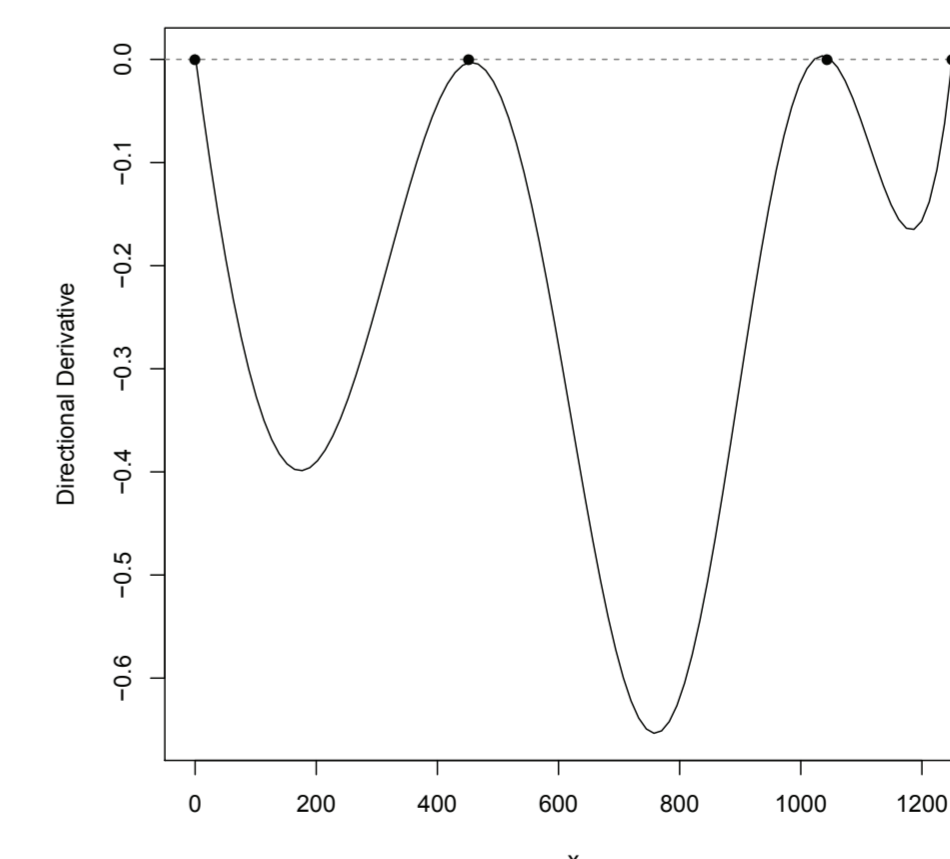


Figure: Equivalence plot confirms the max-min KL-optimality of  $\xi_m$  found by PSO-S-QN for discriminating five toxicology models.

## REFERENCES

- Atkinson, A. C. and Fedorov, V. V. (1975). The design of experiments for discriminating between two rival models. *Biometrika*, 62(1):57–70.
- Chen, R.-B., Chang, S.-P., Wang, W., Tung, H.-C., and Wong, W. K. (2015). Minimax optimal designs via particle swarm optimization methods. *Stat. Comput.*, 25(5):975–988.
- López-Fidalgo, J., Tommasi, C., and Trandafir, P. C. (2007). An optimal experimental design criterion for discriminating between non-normal models. *J. Roy. Stat. Soc. B*, 69(2):231–242.
- Tommasi, C., Martín-Martín, R., and López-Fidalgo, J. (2016). Max-min optimal discriminating designs for several statistical models. *Stat. Comput.*, 26(6):1163–1172.