

# UC San Diego

## UC San Diego Previously Published Works

**Title**

Asynergistic regression based on maximized rank correlation

**Permalink**

<https://escholarship.org/uc/item/8x71t0b6>

**Journal**

Communications in Statistics-Simulation and Computation, 37(1)

**ISSN**

0361-0918

**Authors**

Donohue, Michael

Abramson, Ian

Gamst, Anthony

**Publication Date**

2008

Peer reviewed

# ASYNERGISTIC REGRESSION BASED ON MAXIMIZED RANK CORRELATION

Michael Donohue, Ian Abramson, and Anthony Gamst

Michael Donohue

Dept. of Family and Preventive Medicine

University of California, San Diego

La Jolla, CA 92093

mdonohue@math.ucsd.edu

Ian Abramson

Dept. of Mathematics

University of California, San Diego

La Jolla, CA 92093

Anthony Gamst

Dept. of Family and Preventive Medicine

University of California, San Diego

La Jolla, CA 92093

Key Words: Synergy; Dose-Response; Drug Interactions; Monotonic Index Models.

## ABSTRACT

The property of synergy and its detection are discussed. A response surface is said to possess synergy if it is monotone in each argument and its level curves are convex. Detecting this property is particularly useful in the study of combination drug therapies where the goal is enhanced response or diminished side effect. One way to detect synergy is to fit a surface with linear level curves under the assumption of asynergy and observe the residuals. We explore an algorithm to accomplish this asynergistic regression via a reduction in dimensionality and connections to semiparametric monotonic linear index models. We see that the asynergistic model is a confounded version of the monotonic linear index model where the

linear level curves are not restricted to be parallel.

## 1. INTRODUCTION

For simplicity, we restrict the discussion of synergy to the case of two agents, and to motivate the discussion we will refer to the agents as drug treatments. For example, consider two human immunodeficiency virus (HIV) treatments: the Nucleoside Reverse Transcriptase Inhibitor (NRTI) Emtricitabine (FTC) and the Protease Inhibitor (PI) Atazanavir (ATV). Suppose 300 mg of FTC has the same expected level of increase in the CD4+ cell count as 200 mg of ATV. For some fixed  $\lambda \in (0, 1)$ , the two drugs are synergistic at the combined dose of  $300\lambda$  mg of FTC and  $200(1 - \lambda)$  mg of ATV if the expected response is greater than that of the pure doses. The response at the combined dose is called asynergistic if it is the same as the pure dose, and antagonistic if it is less than the pure dose.

Now to generalize and define these scenarios more rigorously, let  $f(x_1, x_2)$  be the expected dose response to administering amount  $x_1$  of Drug 1 and  $x_2$  of Drug 2. We assume that the response surface is monotone increasing in each argument, and in particular the marginal response curves,  $f_1(x_1) = f(x_1, 0)$  and  $f_2(x_2) = f(0, x_2)$ , are monotone increasing in their respective arguments and onto the same range.

Suppose the expected dose response at a particular dose combination  $(x_1^*, x_2^*)$  is  $y^*$ , i.e.  $f(x_1^*, x_2^*) = y^*$ . Note that  $f_1^{-1}(y^*)$  is the amount of a pure dose of Drug 1 required to attain the response  $y^*$ , and likewise for  $f_2^{-1}(y^*)$  with Drug 2. That is  $y^* = f(f_1^{-1}(y^*), 0) = f(0, f_2^{-1}(y^*))$ .

**Definition 1.** *We say  $f$  is asynergistic, sometimes called additive, dose-additive, zero-interactive, or noninteractive at the dose combination  $(x_1^*, x_2^*)$  if*

$$\frac{x_1^*}{f_1^{-1}(y^*)} + \frac{x_2^*}{f_2^{-1}(y^*)} = 1 \quad (1)$$

*If the “=” in (1) is replaced by “<,”  $f$  is said to be synergistic at  $(x_1^*, x_2^*)$ . Alternatively, if we have “>” instead of “=,”  $f$  is said to be antagonistic at  $(x_1^*, x_2^*)$ .*

The left side of (1) is known as Berenbaum’s Interaction Index (Feng & Kelly, 2004) and the equality is an expression of the collinearity of  $(f_1^{-1}(y^*), 0)$ ,  $(x_1^*, x_2^*)$ , and  $(0, f_2^{-1}(y^*))$ .

The line through  $(f_1^{-1}(y^*), 0)$  and  $(0, f_2^{-1}(y^*))$  is called the  $y^*$ -theoretical line of asynergy (Laska et al., 1994). Asynergy occurs if  $(x_1^*, x_2^*)$  is on its  $y^*$ -theoretical line of asynergy. Synergy occurs if  $(x_1^*, x_2^*)$  is below its  $y^*$ -theoretical line of asynergy. Antagonism occurs if  $(x_1^*, x_2^*)$  is above its  $y^*$ -theoretical line of asynergy. Figure 1 illustrates the three scenarios. The plots include a stylized view of the marginal response curves folded down onto the plane of dose combinations. Each plot represents a different  $y^*$  value and associated  $y^*$ -theoretical line of asynergy.

The methods proposed in this paper primarily deal with assessing the presence of synergy in a global sense. For a given domain,  $D$ , on which  $f$  is defined, we will consider  $f$  to be *asynergistic on  $D$*  if (1) holds for all dose combinations  $(x_1^*, x_2^*)$ . In our context  $D$  will generally be the first quadrant, or some subset. Global synergy and antagonism are defined analogously. In the global context, the three definitions can be classified in terms of the level curves of the response surface. An asynergistic response surface has linear level curves, a synergistic surface has convex level curves, and an antagonistic surface has concave level curves as depicted in Figure 2.

Note that the level curves of an asynergistic surface are not restricted to be parallel, therefore asynergistic surfaces are not limited to planes or even monotone transformations of planes. Asynergistic surfaces are also not necessarily additive in the sense of Hastie & Tibshirani (1986) ( $f(x_1, x_2) = f_1(x_1) + f_2(x_2)$ ). As an example, consider the asynergistic surface  $f(x_1, x_2) = x_1 + (x_1^2 + 4x_2)^{1/2}$ , which produced the linear contours in Figure 2. This surface is asynergistic with a concave marginal curve in  $x_2$  and linear marginal curve in  $x_1$ . In fact, any surface of the form  $f_{a,b}(x_1, x_2) = x_1 + ((x_1 + a)^2 + bx_2)^{1/2}$ ,  $a, b \geq 0$  is asynergistic.

Given a sample  $\{(x_{i1}, x_{i2}, y_i) : i = 1, \dots, n\}$  from the model  $y_i = f(x_{i1}, x_{i2}) + \varepsilon_i$ , our goal is to assess whether  $f$  is asynergistic, synergistic, or antagonistic. In Section 2 we will review existing methods to handle this problem. In Section 3 we will propose our asynergistic regression procedure. In the appendix we discuss consistency.

## 2. REVIEW OF EXISTING METHODS

There are three general classes of approaches to the synergy detection problem and we

will label them with the terminology of Feng & Kelly (2004). The Response Surface Method (RSM) is a classical regression approach along with a synergy assessment procedure. Within the RSM there are parametric and nonparametric approaches in the literature. The Marginal Dose-Response Curve Method (MDRCM) estimates the marginal response curve via data from pure doses of each drug. The most recent contribution to the literature has been regarding the Model-Free Test (MFT) (Feng & Kelly, 2004), which we briefly review.

## 2.1. MODEL-FREE TEST

Laska et al. (1989) introduced the Model-Free Test for synergy which has recently been examined in higher dimensional context by Feng & Kelly (2004). Fix some combination  $(x_1^*, x_2^*)$  and consider for  $r > 0$ , the line through  $(x_1^*, x_2^*)$  with slope  $-r$ :  $(x_2 - x_2^*) = -r(x_1 - x_1^*)$ . Note that  $x_1^* + x_2^*/r$  and  $rx_1^* + x_2^*$  are the respective  $x_1$  and  $x_2$  intercepts. The MFT procedure uses the following fact (Laska et al., 1994; Plummer et al., 1992).

**Theorem 2.** *If there exists an  $r > 0$  such that  $f(x_1^*, x_2^*) > f_1(x_1^* + x_2^*/r)$  and  $f(x_1^*, x_2^*) > f_2(rx_1^* + x_2^*)$ , then  $f$  is synergistic at  $(x_1^*, x_2^*)$ .*

The Model-Free Test begins with an approximation,  $-r$ , such that we believe  $f_1(x_1^* + x_2^*/r) \approx f_2(rx_1^* + x_2^*)$ . (When we have equality in preceding,  $-r$  is the slope of the so-called  *$y^{**}$ -theoretical line of asynergy* through  $(x_1^*, x_2^*)$ , with  $y^{**}$  the expected response at each of the pure doses.) The result is a hypothesis test based on Theorem 2 of the form:

$$H_0 : f(x_1^*, x_2^*) \leq f(x_1^* + x_2^*/r, 0) \text{ or } f(x_1^*, x_2^*) \leq f(0, rx_1^* + x_2^*) \text{ (no synergy)} \quad (2)$$

versus

$$H_1 : f(x_1^*, x_2^*) > f(x_1^* + x_2^*/r, 0) \text{ and } f(x_1^*, x_2^*) > f(0, rx_1^* + x_2^*) \text{ (synergy)}. \quad (3)$$

Laska & Meisner (1989) adapt the test for multiple dose combinations along the same diagonal of slope  $-r$  through the design space.

## 2.2. MARGINAL DOSE-RESPONSE CURVE METHOD

Another approach is to estimate the marginal response curves using data from studies of

responses of pure doses of each drug. These monotone curves are used to generate hypothesis tests for synergy.

Greco et al. (1995) provide a thorough review of the prominent methods developed through 1995. The authors discuss several parametric models for estimating the marginal response curves including the logistic model (Carter et al., 1985) and median effect/sigmoid model (Chou & Talalay, 1983). Kelly & Rice (1990) suggest a hybrid smoothing spline and least squares as a nonparametric alternative. The estimated marginal response curves are generally utilized in one of two ways.

In the first method, one builds a response surface as an extension of the marginal response curve estimates under the assumption of asynergy. That is, a surface with linear level curves is constructed possessing the estimated marginal curves. Kelly & Rice (1990) introduced a searching algorithm for finding these level curves. Once the asynergistic surface is established, an experiment is conducted on combined doses and the observations are compared to the constructed asynergistic surface.

The second methods fits the data to a model with a suitable “synergy parameter,” utilizing the marginal curve estimates. The hypothesis test is then based on the estimation of the synergy parameter. Greco et al. (1995) suggest a parametric model for this approach.

### 2.3. RESPONSE SURFACE METHOD

Finally, the RSM fits the data to a surface all at once, giving no special treatment to the marginal dose-response curves. An assessment is then made as to whether this surface deviates significantly from an asynergistic one.

A slight variation from this general theme, Tibshirani (1990) suggested fitting a bivariate smoother, a gaussian kernel smoother say, and straightening its level curves by a “local principal component line fitting” process. The straightening is accomplished by local (with respect to the observed responses) averaging of dose combinations near the same level curve. Fitting such an asynergistic surface is the focus of the methods we propose. Our methods however, will attempt to fit such a surface directly, that is without first fitting a general surface.

### 3. MAXIMIZING ASYNERGISTIC RANK CORRELATION:

#### THE ASYNERGYMAX ALGORITHM

A special case of an asynergistic surface is one with *parallel* linear level curves. Han (1987), without reference to synergy, introduced a “semiparametric monotonic linear index model” of the form  $Y_i = D \circ F(\mathbf{x}_i' \beta_0, \varepsilon_i)$ . Here  $Y_i$  is an observed response,  $D$  is a monotone increasing function,  $F$  is strictly increasing in both arguments,  $\mathbf{x}_i$  is a known  $p$ -vector,  $\beta_0$  is an unknown  $p$ -vector of interest, and  $\varepsilon_i$  is random error independent of  $\mathbf{x}_i$ . Han also introduced a maximum rank correlation (MRC) estimator

$$\operatorname{argmax}_{\beta \in \mathcal{B}} \frac{1}{n(n-1)} \sum_{i \neq j} \{Y_i > Y_j\} \{\mathbf{x}_i' \beta > \mathbf{x}_j' \beta\} \quad (4)$$

where  $\{\cdot\}$  represents the indicator function and  $\mathcal{B}$  is an appropriate subset of  $\mathcal{R}^p$ , the unit ball say. The necessity of restricting  $\mathcal{R}^p$  becomes apparent in light of the fact that scalar multiples of a particular  $\beta$  yield the same rank correlation. The power of estimators of this type is that they exploit monotonicity without making assumptions about the particular form of  $D$  or  $F$ . Assumptions regarding the error distribution are minimal as well. Sherman (1993) showed (4) is  $n^{1/2}$ -consistent and asymptotically normal.

Cavanagh & Sherman (1998) proposed a class of consistent and asymptotically normal estimators of the form

$$\hat{\beta}_n = \operatorname{argmax}_{\beta \in \mathcal{B}} \sum_{i=1}^n M(Y_i) R_n(\mathbf{x}_i' \beta) \quad (5)$$

where  $R_n$  is the rank function  $R_n(a_i) = \sum_{j=1}^n \{a_j \leq a_i\}$ , for  $(a_1, \dots, a_n) \in \mathcal{R}^n$ , and  $M$  is either deterministic or  $R_n$ . When  $M$  is  $R_n$ , the quantity being maximized in (5) is a linear function of Spearman’s rank correlation coefficient. When  $M = R_n$  we will refer to (5) as the Spearmax estimate of  $\beta_0$ .

One way to approach the generalized asynergistic regression problem is to maximize Spearman’s rank correlation over all asynergistic functions. This amounts to finding the *asynergistically consistent* ordering of the points in the design space which maximizes its rank correlation with the observations. An asynergistically consistent ordering is one associated with a viable asynergistic surface.

Suppose we have an  $n \times 2$  design  $X = (X_1, X_2)$  with associated responses  $Y$ . Let  $\Omega$  be the space of all asynergistic functions on our design space. For each  $f \in \Omega$ , the linear contour through  $(x_{i1}, x_{i2})$  has intercepts  $(\tilde{x}_{fi1}, 0)$  and  $(0, \tilde{x}_{fi2})$ . The vector  $\tilde{X}_{f1}$  determines an ordering of the dose combinations. We seek

$$\tilde{X}_{*1} = \underset{\tilde{X}_{f1}: f \in \Omega}{\operatorname{argmax}} R_n(\tilde{X}_{f1}) R_n(Y) \quad (6)$$

Once  $\tilde{X}_{*1}$  is found, the problem is reduced to simple monotone smoothing accomplished with Ramsay's monotone smoother (Ramsay, 1998) or a constrained B-spline smoother for example. We will describe an algorithm which approximates  $\tilde{X}_{*1}$ , but first it is instructive to discuss known bounds on the target rank correlation.

Recall that the space of asynergistic functions are all functions with linear contours with negative slope. The space of functions with parallel and linear contours with negative slope is clearly contained in  $\Omega$ . We have the lower bound  $\max_{\beta \in \mathcal{B}} R_n(X\beta) R_n(Y) \leq \max_{f \in \Omega} R_n(\tilde{X}_{f1}) R_n(Y)$ , where  $\mathcal{B}$  traces out the first quadrant part of the unit circle (which produces negatively sloped parallel contours).

We invoke the matrix partial ordering:  $(x_1, y_1) \lesssim (x_2, y_2)$  if and only if  $x_1 \leq x_2$  and  $y_1 \leq y_2$ . Clearly any asynergistically consistent ordering must obey the matrix partial ordering. Therefore, consider sorting the design with respect to the partial ordering, breaking ties with the responses,  $Y$ . This ordering is not necessarily asynergistically consistent, but it provides a nontrivial upper bound on the desired rank correlation.

We describe a subroutine which incrementally improves the rank correlation associated with some asynergistically consistent ordering. Take some collection of points which are adjacent in terms of a given asynergistically consistent ordering and place them in the "active bin." Now maximize the total rank correlation over asynergistic functions with parallel contours through the points in the active bin holding the other contours "as fixed as possible." In practice this is done by conducting a grid search through possible contour slopes. For each slope in the grid, if the active contours cross any inactive contours, the slope of the violated inactive contours are also updated. This allows the bin selection to be "fuzzy"



in the sense that points in adjacent bins can be included in the active bin. The resulting rank correlation after all the contours are uncrossed is compared with the previous best asynergistically consistent rank correlation. If the rank correlation is improved the new contours are kept, otherwise we revert to previous contours.

For example, consider a random sample of size 20 as pictured in Figure 3 with the contours resulting from the Spearmax estimate. Rank correlation is 0.8561. The points in the active pool are indicated with “x” instead of “o”. In this case the seventh through tenth points are in the active bin. In Figure 3 we attempt to update the slope through the active pool, but the contours now cross. In Figure 3 we update the inactive contours and compute the rank correlation. If this new rank correlation is greater than 0.8561, we keep the updated contours. If there is no improvement in the rank correlation we revert to the contours in Figure 3. In either case, we then proceed to the next slope value in our grid search. Once all the slope values have been checked, we proceed to the next bin.

In pseudo code, the Asynergymax subroutine begins with current estimates of the intercepts  $\tilde{X}$  such that  $(\tilde{x}_{i1}, 0)$ ,  $(x_{i1}, x_{i2})$ , and  $(0, \tilde{x}_{i2})$  are collinear for all  $i$  and none of these lines cross. We also have an active bin associated with some adjacent collection of intercepts with indices  $a = (a_1, a_2, \dots, a_k)$ .  $U$  is a list of unit vectors,  $u_i$ , which represent a suitably fine grid of directions in  $\mathcal{R}^2$ . Let  $\tilde{X}_a$  denote the matrix of the  $a_1, a_2, \dots, a_k$  rows of  $\tilde{X}$ , and likewise for  $X$ . The subroutine proceeds:

1. Initialize new intercepts  $\tilde{X}^{(1)} := \tilde{X}$ .
2. For each  $u_i$  in  $U$  compute new intercepts  $\tilde{X}_{a1}^{(1)} = X_a(u_{i1}/u_{i2}, 1)'$ , and  $\tilde{X}_{a2}^{(1)} = X_a(1, u_{i2}/u_{i1})'$  associated with lines through the points  $(x_{j1}, x_{j2})$ ,  $j \in a$ , all having normal vector  $u_i$ .
  - (a) For all  $j$  not in  $a$ , if  $R_n(\tilde{x}_{j1}^{(1)}) \neq R_n(\tilde{x}_{j2}^{(1)})$ , give  $(x_{j1}, x_{j2})$  new intercepts as well:  $\tilde{x}_{j1}^{(1)} = x_j'(u_{i1}/u_{i2}, 1)$ , and  $\tilde{x}_{j2}^{(1)} = x_j'(1, u_{i2}/u_{i1})$  until  $R_n(\tilde{X}_1^{(1)}) = R_n(\tilde{X}_2^{(1)})$  (ensuring no lines cross).
  - (b) Compute the (unscaled) rank correlation associated with new intercepts and record as  $r_i = R_n(Y)'R_n(\tilde{X}^{(1)})$ .

- (c) Reset intercepts  $\tilde{X}^{(1)} = \tilde{X}$  and repeat until all grid values have been evaluated.
- 3. Finally update the intercepts by those associated with greatest rank correlation. That is, once again compute the intercepts for  $u_i$  where  $r_i$  is the maximal component of the list  $r$ .
- 4. Set  $\tilde{X} := \tilde{X}^{(1)}$ , the intercepts which maximize rank correlation with respect to the current active bin and proceed to next active bin.

We get good results by allowing the bin sizes to follow the geometric progression  $(n, n/2, n/4, n/8, \dots, 1)$ , but again, these bin sizes are allowed to expand as necessary with the “uncrossing” step. The first iteration, with one bin of size  $n$ , is simply the Spearman estimate. The next iteration takes the first  $n/2$  points in the first bin and the rest in the second bin, and so on. Each iteration divides the previous bins in half until there is only one point in each of  $n$  bins.

Note the apparently higher variability in the lower left. The estimates in the lower left and upper right are susceptible to this defect given the experiment’s square sampling design. There are fewer observations on which to base the estimate of the slope through points in these regions. One could fix a larger bin size over regions of this type to reduce this type of estimation variance. For example, if the bin size in the lower left was not allowed to be less than 4, the first four contours would necessarily have the same slope.

#### 4. SIMULATIONS AND A TEST OF GLOBAL SYNERGY

In this experiment we used a random design with 50 dose combinations with each dose sampled independently from the  $\text{Unif}(0,1)$  distribution. We approximated the power of the test using 1000 trials with the synergistic surfaces  $f_\delta(x_1, x_2) = x_1 + (x_1^2 + 4x_2)^{1/2} + x_1x_2\delta$ ,  $\delta$  ranging from 0 (asynergy) to 8 (strong synergy). Two procedures were used. First, Asynergymax was applied to estimate the contours, then a constrained B-spline smoother monotone smoothed the responses. Second, Spearman was applied followed by the constrained B-spline smoother.

Global synergy was assessed using the asynergistic fit from the algorithms by examining a residual plot of,  $\{y_i - \hat{f}(x_{1i}, x_{2i})\}$ , versus  $\{\tan^{-1}(x_{2i}/x_{1i})\}$ . This can be viewed as a global version of the MFT performed on all of the interior observations at once.

Figure 4 illustrates these plots from simulations using various  $\delta$  values. This plot detects patterns in the residuals with respect to location relative to the axes. In the presence of global synergy, the asynergistic fit will overestimate the response for doses near each margin and underestimate response from doses in the interior of the grid. Therefore if we fit a quadratic to these residuals using an ordinary least squares regression, we would expect to see a negative quadratic term,  $\theta$  say, in the presence of synergy. Alternatively, if the true model is asynergistic, the quadratic term should be near zero. If the truth is antagonism, we expect a positive quadratic term. This suggests a one-sided hypothesis test for synergy of the form:  $H_0 : \theta = 0$  (asynergistic) versus  $H_1 : \theta < 0$  (synergy), where  $\theta$  is the estimated quadratic term from the fit of the residuals. A  $t$ -test was applied to this estimate with significance level  $\alpha = 0.1$ . To approximate the power of this test, we ran the algorithm 1000 times for each  $\delta$ .

As a comparison, we also estimated the power of the test when the true asynergistic contours are known and when the true marginal dose-response curves (MDRCs) are known. When the true contours are known, we still apply a monotone smoother to the responses based on this known ordering. When the MDRCs are known, we base our test on the residuals from the surface  $f_0$ , i.e. when  $\delta = 0$ . Figure 5 and Table 1 depict the results. Figure 5 adjusts the approximated power curves so the intercepts (the significance level) agree for each curve. For this family of synergistic surfaces and this test for synergy, both procedures perform well and are as good as knowing the true contours.

In this experiment the power curve is drastically improved when we know the true MDRCs. This is due to the fact that responses from high concentrations of both drugs are compared to responses from extreme pure doses which we do not observe. Given the information on  $D = [0, 1] \times [0, 1]$ , we cannot estimate responses from a pure dose of either drug greater than 1. When  $\delta$  is large, responses from these pure doses greater than 1 are

small relative to responses from dose combination where both concentrations are near 1. The residuals based on the MDRCs are therefore large from doses on the upper right triangle which results in greater convexity in the residual plot and more synergy conclusions.

The experiment was repeated with one adjustment. Instead of taking dose combinations,  $(x_{1i}, x_{2i})$ , uniform on the unit square, we take  $(x_{1i}, x_{2i})$  uniform on a *triangular* region,

$$D = \{(x_1, x_2) : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq x_1\}.$$

The results of the simulation are depicted in Figure 6 and adjusted so that significance levels agree in Table 1. Again we see that both procedures detect synergy well. In this case the procedures perform as well as knowing either the true contours or the true MDRCs.

Since the Spearmax class of surfaces (parallel linear contours) is contained in the Asynergymax class (linear contours), the Asynergymax procedure must match or exceed Spearmax in terms of mean squared error (MSE). When the alternative is true and contours are convex, both procedures produce biased estimates of the quadratic term of the fit through the residual plot. In the “square design” experiment, the bias from the Spearmax procedure tends to favor a synergy conclusion more so than the Asynergymax. The opposite was true in the “triangular design” experiment. When the null hypothesis is true and the contours are in fact linear, this bias will go away asymptotically using the Asynergymax procedure, but in general it will remain using the Spearmax procedure. Therefore, one advantage of using Asynergymax is that false positive tests for synergy (Type I errors) could be reduced if the Spearmax bias favors a synergy conclusion.

Another observation is that our random design experiments offer little information regarding contours through the upper-right-hand or lower-left-hand corner of the unit square. There are relatively fewer observations near these contours, and therefore greater variance in estimating the slope, say, of these contours. The “triangular design” reduces this variances by eliminating one of these extreme regions. Sampling from a triangular region is also more realistic, since in practice a large dose of multiple agents would most likely be avoided.

One obvious drawback of a global approach to synergy detection is that the true level of synergistic effect is not necessarily consistent throughout the region of interest. For example,

for the curves used in our experiments the contours are most convex in the lower left region of the design. In the upper right region, the contours are nearly linear. This was another motivation for sampling from the triangular region, where the true contours were most convex and nonparallel.

Another issue to consider is whether the contours are parallel or “splayed” (nonparallel). In particular, if the underlying surface is synergistic but the contours are parallel, the advantage of Asynergymax over Spearmax is lost. For example, a surface with concentric parallel contours such as the cone  $2 - (x_1 - 1)^2 - (x_2 - 1)^2$  would yield similar results from both procedures since there is no splaying effect.

As discussed in the MFT literature (Laska & Meisner 1989), we can detect synergy well even when the slope of the  $y^{**}$ -theoretical line of asynergy is not accurately estimated. In fact optimal power is not guaranteed to accompany the true slope of this line. Similarly, the Spearmax procedure can yield good power with our proposed global test even though Asynergymax is better at estimating the slopes of the  $y^{**}$ -theoretical lines of asynergy.

## 5. APPENDIX

We can adapt the methods of proving Spearmax consistency as developed by Han (1987), and Cavanagh & Sherman (1998) to the purpose of proving consistency of the Asynergymax procedure. First we describe how asynergistic functions on  $\mathcal{R}_+^p$  can be identified with  $p$  monotone increasing functions.

Let  $f$  be an asynergistic function on  $\mathcal{R}_+^p$ . Any asynergistic model of the form

$$Y_i = f(\mathbf{x}_i) + \varepsilon_i, \tag{7}$$

where  $\varepsilon_i$  are i.i.d. error terms, can be written in terms of  $p$  monotone increasing functions on  $\mathcal{R}_+$ . Given a set of  $p - 1$  monotone increasing functions  $\alpha_i : \mathcal{R}_+ \rightarrow \mathcal{R}_+$ , onto  $\mathcal{R}_+$ , let

$\psi(\mathbf{x}; \alpha_1, \dots, \alpha_{p-1})$  be the number  $x$  such that  $\mathbf{x}$  is on the hyperplane through the points

$$\begin{aligned} &(x, 0, \dots, 0), \\ &(0, \alpha_1(x), 0, \dots, 0), \\ &\quad \vdots \\ &(0, \dots, 0, \alpha_{p-2}(x), 0), \text{ and} \\ &(0, \dots, 0, \alpha_{p-1}(x)). \end{aligned} \tag{8}$$

Then for the appropriate monotone increasing functions,  $F, \alpha_{01}, \dots, \alpha_{0(p-1)}$ , the model (7) can be written  $Y_i = F(\psi(\mathbf{x}_i; \alpha_{01}, \dots, \alpha_{0(p-1)})) + \varepsilon_i$ . Each level set of the underlying asynergistic  $f$  is a hyperplane through the points

$$\begin{aligned} &(x, 0, \dots, 0), \\ &(0, \alpha_{01}(x), 0, \dots, 0), \\ &\quad \vdots \\ &(0, \dots, 0, \alpha_{0p-2}(x), 0), \text{ and} \\ &(0, \dots, 0, \alpha_{0p-1}(x)), \end{aligned}$$

for some  $x \in \mathcal{R}_+$ . Note that given the asynergistic function  $f$  and its first marginal dose-response curve  $f_1(x) = f(x, 0, \dots, 0)$ , we can write

$$\psi(\mathbf{x}; \alpha_{01}, \dots, \alpha_{0(p-1)}) = f_1^{-1}(f(\mathbf{x})).$$

That is,  $\psi$  maps points in  $\mathcal{R}_+^p$  to the first axis along the level sets of  $f$ . Under the assumption  $\alpha_i$  is onto  $\mathcal{R}_+$  for all  $i$ , one can always determine  $\psi(\mathbf{x}; \alpha_1, \dots, \alpha_{p-1})$  numerically via a searching algorithm: increase  $x$  until the hyperplane defined by (8) contains  $\mathbf{x}$ . An analytic formulation of  $\psi$  given the  $\alpha_i$  is also possible if one can solve the equation of the hyperplane through the points (8) and  $\mathbf{x}$  for  $x$ .

For any asynergistic function,  $f$ , there exists such a unique set of monotone functions  $F, \alpha_{01}, \dots, \alpha_{0(p-1)}$ . In fact, we can write the functions explicitly in terms of  $f$  and marginal dose-response curves,  $f_i$ , as

$$\alpha_{0i}(x) = f_{i+1}^{-1} \circ f(x, 0, \dots, 0) = f_{i+1}^{-1} \circ f_1(x)$$

and  $F(x) = f_1(x)$ .

The  $\alpha_{0i}$  functions are compositions of monotone functions by definition of asynergy, and therefore monotone themselves as is  $F$ . However, it is important to observe that distinct asynergistic functions may have the same  $\alpha_{01}, \dots, \alpha_{0(p-1)}$  (with a different  $F$ ).

For the consistency argument, we are forced to restrict the class of allowable  $\alpha_i$  via a finite basis. Fix some collection of  $d$  monotone increasing basis functions  $\nu_i : \mathcal{R}_+ \rightarrow \mathcal{R}_+$ ,  $i = 1, \dots, d$ . Consider the class of monotone increasing functions

$$\{\alpha : \mathcal{R}_+ \rightarrow \mathcal{R}_+ : \alpha = \beta_1 \nu_1 + \dots + \beta_d \nu_d, \beta \in \mathcal{R}_+^d\}. \quad (9)$$

Now for  $\beta = (\beta_{11}, \dots, \beta_{1d}, \dots, \beta_{(p-1)1}, \dots, \beta_{(p-1)d})' \in \mathcal{R}_+^{d(p-1)}$  we write

$$\psi(\mathbf{x}; \beta) = \psi(\mathbf{x}; \alpha_1, \dots, \alpha_{p-1})$$

where  $\alpha_i = \beta_{i1} \nu_1 + \dots + \beta_{id} \nu_d$ .

Let  $\mathcal{B}$  be some compact subset of  $\mathcal{R}_+^{d(p-1)}$ . We assume observations  $\{(\mathbf{x}_i, Y_i) : \mathbf{x}_i \in \mathcal{R}^p, Y_i \in \mathcal{R}, i = 1, \dots, n\}$  are generated from the model

$$Y_i = F(\psi(\mathbf{x}_i; \beta_0)) + \varepsilon_i, \beta_0 \in \mathcal{B}$$

for some unknown monotone  $F$ .

We prove Asynergymax is consistent for estimating  $\beta_0$  in a manner analogous to Cavanagh & Sherman's (1998) proof that Spearmax is consistent. As in the Spearmax setting we are not particularly concerned with the exact form of the monotone  $F$ , but we wish to exploit its monotonicity as a function of its level sets to estimate the “form” of those level sets.

Let  $M : \mathcal{R} \rightarrow \mathcal{R}_+$  be a deterministic increasing function. Define

$$G_n(\beta) = \frac{1}{n(n-1)} \sum_{i=1}^n M(Y_i) R_n(\psi(\mathbf{x}_i; \beta))$$

or equivalently

$$G_n(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} M(Y_i) \{\psi(\mathbf{x}_i; \beta) > \psi(\mathbf{x}_j; \beta)\}$$

and  $\beta_n = \operatorname{argmax}_{\beta \in \mathcal{B}} G_n(\beta)$ . As in Cavanagh & Sherman (1998), we first prove consistency of  $\beta_n$ , then replace the function  $M$  with the  $R_n(\cdot)$  function and show the argument remains valid. Essentially, everywhere the literature refers to  $\mathbf{x}'\beta$ , we replace it with  $\psi(\mathbf{x}; \beta)$  and demonstrate that the proof still holds.

We will require the following assumptions:

- (A0)  $\mathbf{E}[M(Y)|\mathbf{x}]$  depends on  $\mathbf{x}$  only through  $\psi(\mathbf{x}; \beta_0)$ .
- (A1)  $\mathbf{E}[M(Y)|\mathbf{x}]$  is monotone increasing function of  $\psi(\mathbf{x}; \beta_0)$ .
- (A2) The support of  $\mathbf{x}$  is not contained in a linear subspace of  $\mathcal{R}^p$ .
- (A3)  $\mathbf{x}$  has a positive Lebesgue density on  $\mathcal{R}_+^p$ .
- (A4)  $\mathcal{B}$  is a compact subset of  $\mathcal{R}_+^{d(p-1)}$ .
- (A5)  $\mathbf{E}[M(Y)]^2 < \infty$ .
- (A6)  $M : \mathcal{R} \rightarrow \mathcal{R}_+$  and  $M$  is increasing.

Assumptions (A0)-(A4) are required to ensure identifiability. Assumption (A5) allows us to invoke a zero-mean U-process result of Sherman (1994). The assumption that  $M$  is positive valued (A6) is not included in the analogous Spearman proof, but it shortens our proof and is still broad enough to encompass the  $R_n(\cdot)$  rank function which is bounded below by 1.

Let  $G(\beta) = \mathbf{E}M(Y_1)\{\psi(\mathbf{x}_1; \beta) > \psi(\mathbf{x}_2; \beta)\} = \mathbf{E}G_n(\beta)$ .

Lemma 3.  $G(\beta)$  is uniquely maximized at  $\beta_0$ .

*Proof.* Using assumption (A0), we let  $H(\psi(\mathbf{x}; \beta_0)) = \mathbf{E}[M(Y)|\mathbf{x}]$ . We have

$$\begin{aligned}
G(\beta) &= \frac{1}{2}(\mathbf{E}M(Y_1)\{\psi(\mathbf{x}_1; \beta) > \psi(\mathbf{x}_2; \beta)\} \\
&\quad + \mathbf{E}M(Y_2)\{\psi(\mathbf{x}_1; \beta) < \psi(\mathbf{x}_2; \beta)\}) \\
&= \frac{1}{2}\mathbf{E}(H(\psi(\mathbf{x}_1; \beta_0))\{\psi(\mathbf{x}_1; \beta) > \psi(\mathbf{x}_2; \beta)\} \\
&\quad + H(\psi(\mathbf{x}_2; \beta_0))\{\psi(\mathbf{x}_1; \beta) < \psi(\mathbf{x}_2; \beta)\})
\end{aligned} \tag{10}$$



$H$  must be a monotone increasing function of  $\psi(\mathbf{x}; \beta_0)$  by (A1). Therefore, if  $\beta = \beta_0$ , (A2) ensures the indicators in the above almost surely “choose” the larger of  $H(\psi(\mathbf{x}_1; \beta_0))$  and  $H(\psi(\mathbf{x}_2; \beta_0))$ . As a result,

$$G(\beta_0) = \frac{1}{2} \mathbf{E} \max(H(\psi(\mathbf{x}_1; \beta_0)), H(\psi(\mathbf{x}_2; \beta_0))).$$

and for any  $\beta \in \mathcal{B}$ ,  $G(\beta) \leq G(\beta_0)$ .

Now we show  $\beta_0$  uniquely maximizes  $G(\beta)$ . Suppose for some  $\beta \in \mathcal{B}$

$$G(\beta) = G(\beta_0) = \frac{1}{2} \mathbf{E} \max(H(\psi(\mathbf{x}_1; \beta_0)), H(\psi(\mathbf{x}_2; \beta_0))). \quad (11)$$

We claim there exists sets  $A_1, A_2 \subset \mathcal{R}_+^p$  such that  $\mathbf{x}_1 \in A_1$  and  $\mathbf{x}_2 \in A_2$  implies

$$\psi(\mathbf{x}_1; \beta_0) > \psi(\mathbf{x}_2; \beta_0), \text{ but } \psi(\mathbf{x}_1; \beta) < \psi(\mathbf{x}_2; \beta).$$

We show such sets exist in the case  $p = 2$ . The higher dimensional argument is analogous. If  $\beta_0 \neq \beta$ , there must be a level set associated with  $\beta_0$  which intersects with a level set associated with  $\beta$ . When  $p = 2$ , these level sets are lines which intersect at a point. Let  $\alpha_0$  be the monotone function associated with the parameter  $\beta_0$  and let  $\alpha$  be the monotone function associated with  $\beta$ .

We refer to the notation in Figure 7 where we labeled one of the wedges adjacent to the intersection of the level sets  $A_1$  and another  $A_2$ . Now for  $\mathbf{x}_1 \in A_1$  and  $\mathbf{x}_2 \in A_2$ , by the monotonicity of  $\alpha_0$  and  $\alpha$  we have

$$\psi(\mathbf{x}_1; \beta_0) > \psi(\mathbf{x}_2; \beta_0) \text{ and } \psi(\mathbf{x}_1; \beta) < \psi(\mathbf{x}_2; \beta)$$

as desired.

Now we invoke the assumption that  $M$  is positive valued and  $\mathbf{x}$  has a positive Lebesgue

density on  $\mathcal{R}_+^p$ . We have

$$\begin{aligned}
G(\beta_0) - G(\beta) &= \mathbf{E}M(Y_1)\{\psi(\mathbf{x}_1; \beta_0) > \psi(\mathbf{x}_2; \beta_0)\} \\
&\quad - \mathbf{E}M(Y_1)\{\psi(\mathbf{x}_1; \beta) > \psi(\mathbf{x}_2; \beta)\} \\
&\geq \mathbf{E}\big(M(Y_1)\big(\{\psi(\mathbf{x}_1; \beta_0) > \psi(\mathbf{x}_2; \beta_0)\} \\
&\quad - \{\psi(\mathbf{x}_1; \beta) > \psi(\mathbf{x}_2; \beta)\}\big)\{\mathbf{x}_1 \in A_1\}\{\mathbf{x}_2 \in A_2\}\big) \\
&= \mathbf{E}\big(M(Y_1)\{\mathbf{x}_1 \in A_1\}\big)\mathbf{P}(\mathbf{x}_2 \in A_2) \\
&= \mathbf{E}\big(M(Y_1)\{\mathbf{x}_1 \in A_1\}\big)\mathbf{P}(\mathbf{x}_2 \in A_2) \\
&> 0.
\end{aligned}$$

Therefore  $\beta_0$  is a unique maximizer of  $G$ . □

The consistency arguments of Sherman (1994) and Cavanagh & Sherman (1998) now follow with minor modifications.

## 6. TABLES AND FIGURES

Table 1: Approximate power using Spearmax and Asynergymax

$\delta$	Square Design		Triangular Design	
	Spearmax	Asynergymax	Spearmax	Asynergymax
0	0.107	0.092	0.088	0.102
2	0.212	0.191	0.196	0.207
4	0.362	0.354	0.376	0.391
6	0.547	0.540	0.569	0.593
8	0.721	0.691	0.727	0.735

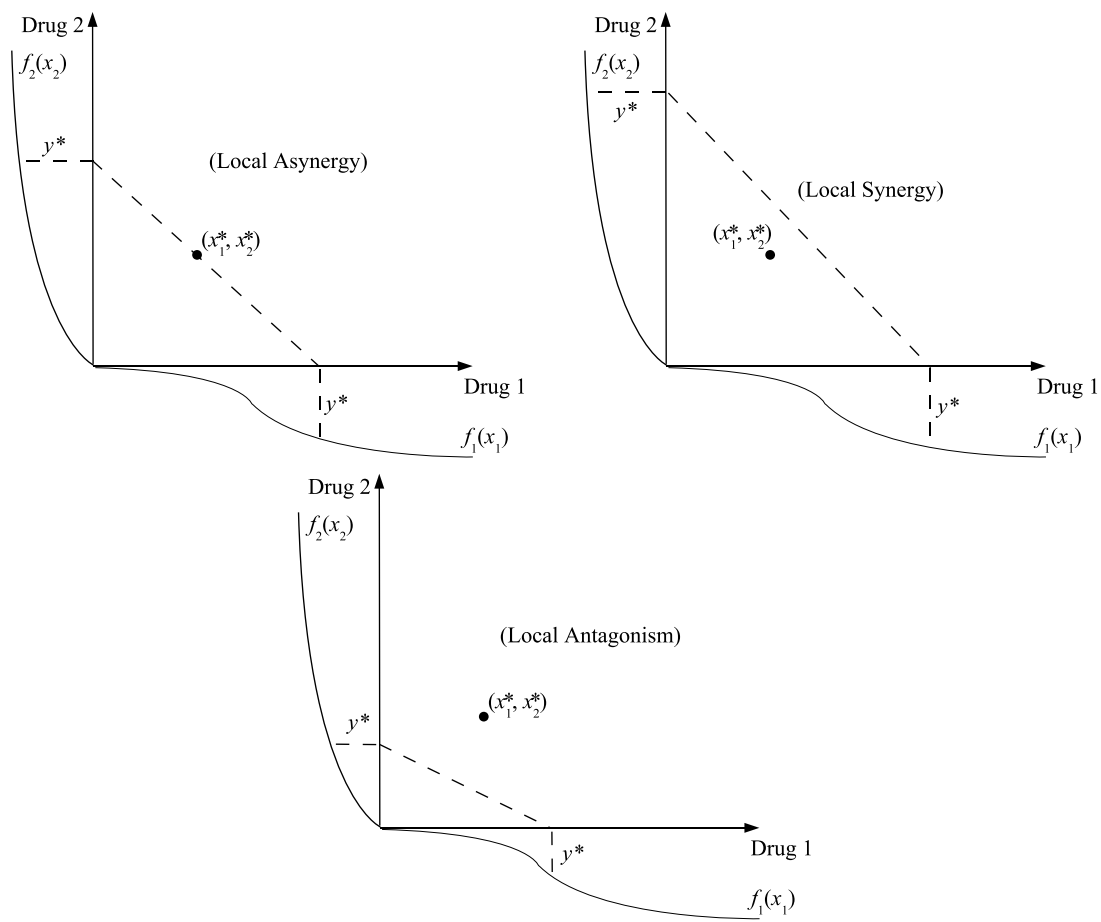


Figure 1: Local asynergy, synergy, and antagonism.

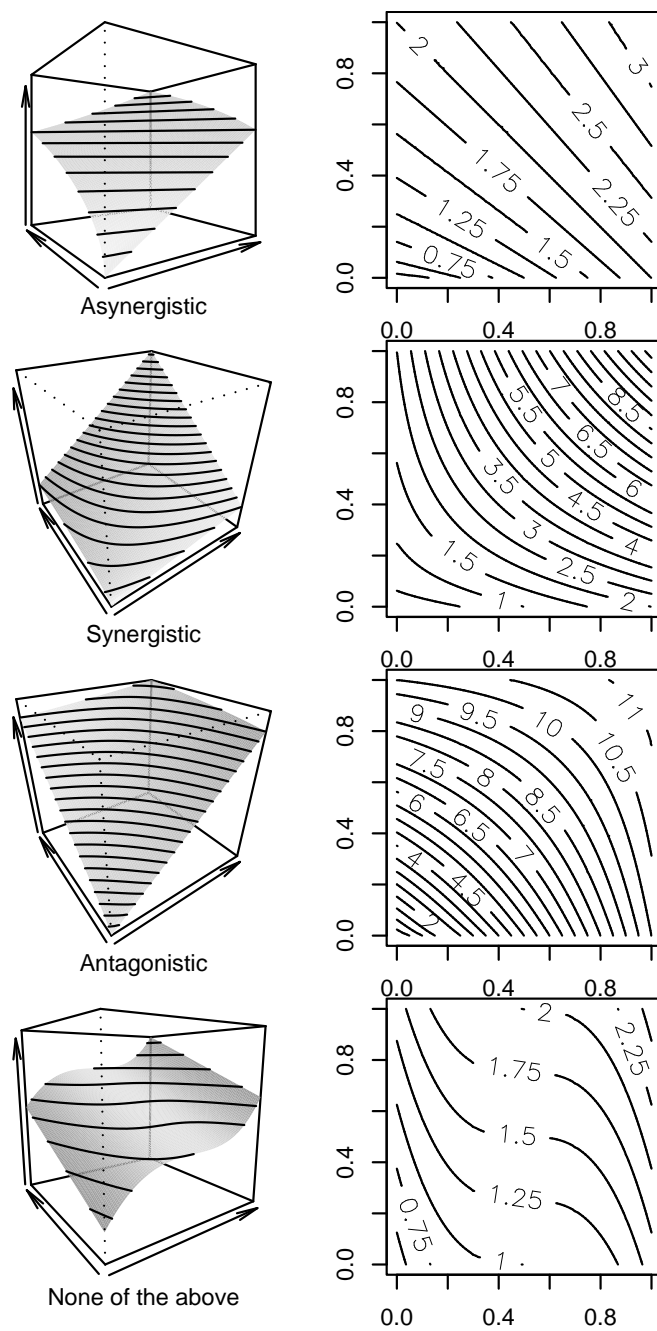


Figure 2: Global asynergy, synergy, and antagonism surfaces and contours.

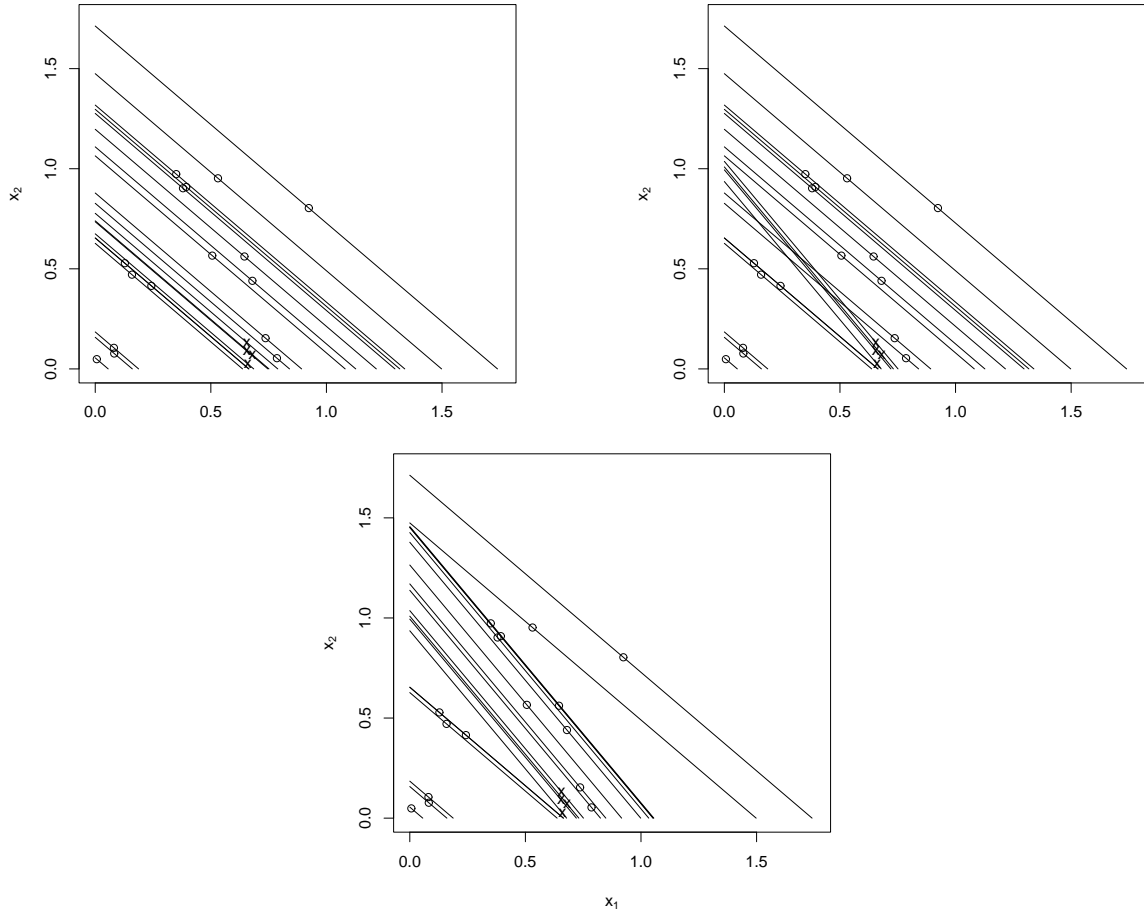


Figure 3: The Asynergymax subroutine.

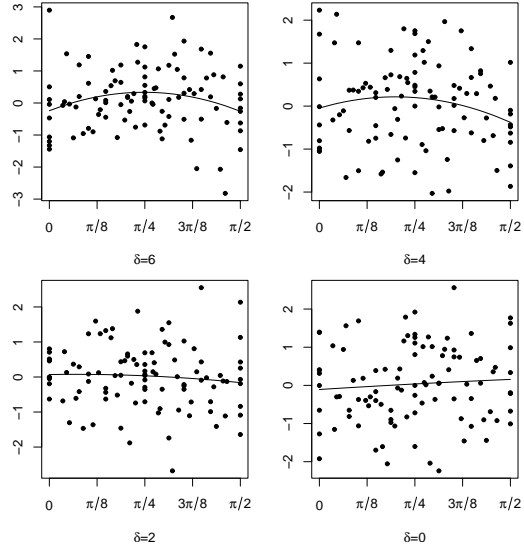


Figure 4: Residuals plotted against  $\tan^{-1}(x_2/x_1)$ .

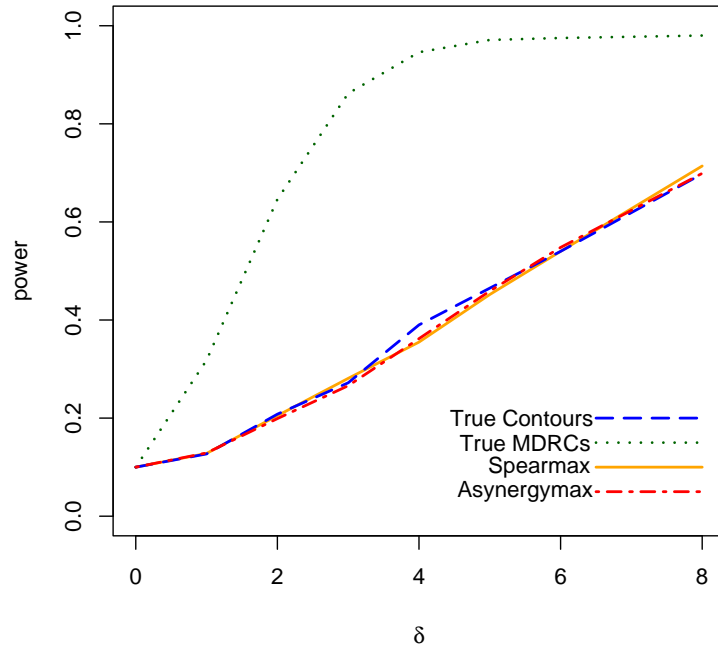


Figure 5: Approximate power curves using Spearmax and Asynergymax (square design space).



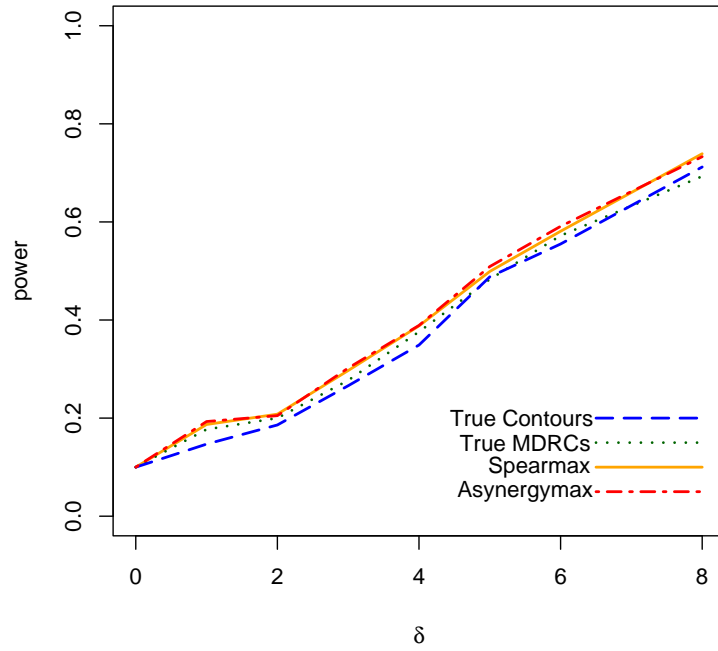


Figure 6: Approximate power curves using Spearmax and Asynergymax (triangular design space).

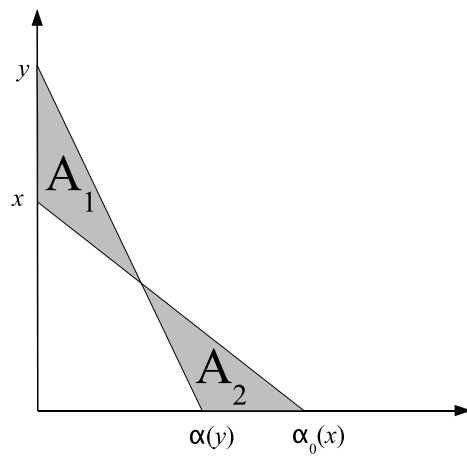


Figure 7: Intersecting contours.

## BIBLIOGRAPHY

- Carter, W. H., Jones, D. E., & Carchman, R. A. (1985), "Application of response surface methods for evaluating the interactions of soman, atropine, and pralidoxime chloride," *Fundamental and Applied Toxicology*, 5, S232–S241.
- Cavanagh, C., & Sherman P. (1998), "Rank estimators for monotonic index models," *J. Econometrics*, 84, 351–381.
- Chou, T. C., & Talalay, P. (1983), "Analysis of combined drug effects: A new look at a very old problem," *Trends in Pharmacological Sciences*, 4, 450–454.
- Feng, P., & Kelly, C. (2004), "An extension of the model free test to test synergy in multiple drug combinations," *Biom. J.*, 46, 293–304.
- Greco, W. R., Bravo, G., Parsons, J. C. (1995), "The search for synergy: A critical review from a response surface perspective," *Pharmacological Reviews*, 47, 331–382.
- Han, A. K. (1987), "Nonparametric analysis of a generalized regression model. The maximum rank correlation estimator," *J. Econometrics*, 35, 303–316.
- Hastie, T., & Tibshirani, R. (1986), "Generalized additive models," *Statistical Science*, 1, 297–309.
- Kelly, C., & Rice, J. (1990), "Monotone smoothing with application to dose-response curves and the assessment of synergism," *Biometrics*, 46, 1071–1085.
- Laska, E. M., & Meisner, M. J. (1989), "Testing whether an identified treatment is best," *Biometrics*, 45, 1139–1151.
- Laska, E. M., Meisner, M., & Siegel, C. (1994), "Simple designs and model-free tests for synergy," *Biometrics*, 50, 834–841.
- Plummer, J. L., Cmielewski, P. L., Gourlay, G. K., Owen, H., & Cousins, M. J. (1992), "Antinociceptive and motor effects of intrathecal morphine combined with intrathecal cloni-

dine, noradrenaline, carbachol or midazolam in rats,” *Pain*, 49, 145–52.

Ramsay, J. O. (1998), “Estimating smooth monotone functions,” *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 60, 365–375.

Sherman, R. P. (1993), “The limiting distribution of the maximum rank correlation estimator,” *Econometrica*, 61, 123–137.

Sherman, R. P. (1994), “Maximal inequalities for degenerate  $U$ -processes with applications to optimization estimators,” *Ann. Statist.*, 22, 439–459.