Copula-based robust optimal block designs

W.G. Müller\textsuperscript{a} A. Rappold\textsuperscript{b} and D.C. Woods\textsuperscript{c}

\textsuperscript{a} Department of Applied Statistics, Johannes Kepler University Linz, Austria
\textsuperscript{b} Plasser & Theurer Connected G.m.b.H, Hagenberg im Mühlkreis, Austria
\textsuperscript{c} Southampton Statistical Sciences Research Institute, University of Southampton, UK

Blocking is often used to reduce known variability in designed experiments by collecting together homogeneous experimental units. A common modelling assumption for such experiments is that responses from units within a block are dependent. Accounting for such dependencies in both the design of the experiment and the modelling of the resulting data when the response is not normally distributed can be challenging, particularly in terms of the computation required to find an optimal design. The application of copulas and marginal modelling provides a computationally efficient approach for estimating population-average treatment effects. Motivated by an experiment from materials testing, we develop and demonstrate designs with blocks of size two using copula models. Such designs are also important in applications ranging from microarray experiments to experiments on human eyes or limbs with naturally occurring blocks of size two. We present methodology for design selection, make comparisons to existing approaches in the literature and assess the robustness of the designs to modelling assumptions.

Key words: Binary response; pseudo-Bayesian D-optimality; equivalence theorem; generalized linear model; marginal model.

1 Introduction and motivation

Statistical design of experiments underpins much quantitative work in the biological, physical and engineering sciences, providing a principled approach to the efficient allocation of (typically sparse) experimental resources to address the aims of the study. Often, experiments aim to understand a process by modeling discrete data, for example arising from the observation of a binary or count response. For completely randomized experiments, assuming homogeneous experimental units, a generalized linear model (GLM) may provide an appropriate description and there has been much research into the construction of optimal and efficient designs for multi-factor GLMs, including [Woods et al. (2006), Dror and Steinberg (2006, 2008) and Russell et al. (2009)]. See Atkinson and Woods (2015) for a comprehensive review.

When heterogeneous experimental units can be grouped into more homogenous groups, or blocks, accounting for this grouping can improve the precision of inferences made from the experimental data. Methods to find block designs for discrete data have recently been proposed by, amongst others, [Woods and van de Ven (2011), Niaparast and Schwabe (2013) and Waite and Woods (2015)]. Two modeling paradigms have been adopted in the design literature: conditional models where the joint distribution of the data is derived by explicitly including block-specific random effects (e.g. generalized linear mixed models, Breslow and Clayton (1993)); and marginal models, where the dependence structure of the data is specified separately from the marginal distribution of each response (e.g. with parameters estimated via generalized estimating equations (GEEs), Liang and Zeger (1986)). For the linear model, these two modeling approaches coincide. In this paper, we find optimal designs under a marginal...
modeling approach when the intra-block dependence structure is defined via a copula. Such models are particularly appropriate when block effects are not of interest in themselves and the aim of the experiment is to understand the effects of treatment factors averaged across blocks. Optimal designs for marginal models using alternative definitions of the dependence structure have been found by Hughes-Oliver (1998), Atkinson and Ucinski (2004) and van de Ven and Woods (2014).

Although our methods can be generalized to arbitrary block sizes, we focus on the important special case of experiments with blocks of size two (see Godolphin, 2018). Such blocks occur routinely in microarray experiments (Bailey, 2007, Kerr, 2012) and in experiments on people, for example with eyes or arms as experimental units (David and Kempton, 1996). Practical motivation for our work comes from a materials science experiment. In Section 3 we find designs appropriate for aerospace materials testing experiments similar to those performed by our collaborators at the UK Defence Science and Technology Laboratory. The aim of these experiments is to compare the thermal properties of a set of novel materials against a reference material. In particular, one aim is to assess the probability of failure due to the exposure to extreme (high) temperatures. The experiment is performed using a arc jet to heat material samples which are held in one of six “wedges”, each of which holds a pair of samples on a strut attached to a circular carousel, see Figure 1. Hence, the experiment can be considered as a block design with six blocks, each containing two units. In the particular experiment considered here, six materials were tested, a reference and five novel samples. A variety of measures are made on each tested sample, including a visual inspection of quality to assess material failure which leads to a binary (pass/fail) response. It is this response for which we find optimal designs.

Figure 1: Arc jet carousel, struts and “wedges” (left) and schematic (right). In addition to the six wedges for holding material samples, the carousel had two further wedges used for temperature measurement.

In common with most nonlinear models, the performance of a given design for a copula-based GLM model may depend on the values of the model parameters that define both the marginal model and the dependence structure. If strong prior information is available, then locally optimal designs can be sought for given values of the model parameters. Otherwise, Bayesian (e.g. Overstall and Woods, 2017) or maximin (e.g. King and Wong, 2000) approaches can be adopted. In common with much of the recent literature on designs for GLMs, we find optimal designs robust to the values of the model parameters via a pseudo-Bayesian approach (e.g. Atkinson et al., 2007, ch. 18), with a classical quantity for design performance averaged
with respect to a prior distribution on the parameters. Here, we adopt variants of $D$-optimality for design selection.

The remainder of the paper is organized as follows. In Section 2 we introduce the statistical models we employ, including copulas, and develop design methods for blocked experiments. An illustrative comparison is made to previous design approaches based on GEEs using an example from Woods and van de Ven (2011). In Section 3 we demonstrate and assess our methods via application to the materials testing example. In particular, we show how prior information on the parameters influences the choice of optimal design. We provide a brief discussion and some areas for future work in Section 4.

2 Designs for copula-based marginal models

Suppose the experiment varies $m$ treatment factors, $\mathbf{x}_T = (x_1, \ldots, x_m)$, and the experiment has $b$ blocks of size $k$; throughout, our examples will assume $k = 2$. The $j$th unit in the $i$th block receives treatment $\mathbf{x}_{ij} = (x_{1ij}, \ldots, x_{mij})$ ($i = 1, \ldots, b; j = 1, \ldots, k$) and realizes observation $Y_{ij}$. The $x_{ij}$ are chosen from a standardized design space $\mathcal{X} = [-1, 1]^m$ and are not necessarily distinct. Independence of observations $Y_{ij}, Y_{i'j'}$, for $i, i' = 1, \ldots, b; j, j' = 1, \ldots, k$, is assumed across blocks ($i \neq i'$) but we allow dependence within a block ($i = i'$), which we describe via a copula model.

2.1 Statistical modeling via copulas

The problem of specifying a probability model for dependent random variables $Y_{i1}, \ldots, Y_{jk}$ can be simplified by expressing the corresponding $k$-dimensional joint distribution $F_{Y_{i1}, \ldots, Y_{ik}}$ in terms of marginal distributions $F_{Y_{i1}}, \ldots, F_{Y_{ik}}$, and an associated $k$-copula (or dependence function) $C$ defined as follows (cf. Nelsen, 2007).

**Definition 2.1.** A $k$-copula is a function $C : [0, 1]^k \to [0, 1], k \geq 2$, with the following properties:

1. (uniform margins) for every $u \in [0, 1]^k$, if at least one coordinate of $u$ is 0, then $C(u) = 0$,

   and if all coordinates of $u$ are 1 except $u_i$, then $C(u) = u_i$.

2. (k-increasing) for all $a, b \in [0, 1]^k$ such that $a \leq b$,

   $V_C([a, b]) \geq 0$,

where $V_C$ is the measure induced by $C$ on $[0, 1]^k$.

The connection between a copula and a joint probability distribution is given by Sklar’s Theorem (Sklar, 1959), which affirms that for every $k$-dimensional joint distribution $F_{Y_{i1}, \ldots, Y_{ik}}$ with marginal distributions $F_{Y_{i1}}, \ldots, F_{Y_{ik}}$, there exists a $k$-copula $C$, defined as in Definition 2.1, such that

$$F_{Y_{i1}, \ldots, Y_{ik}}(y_1, \ldots, y_k) = C(F_{Y_{i1}}(y_1), \ldots, F_{Y_{ik}}(y_k)),$$

(1)

for all $y_1, \ldots, y_k \in \mathbb{R}$. Conversely, if $C$ is a $k$-copula and $F_{Y_{i1}}, \ldots, F_{Y_{ik}}$ are distribution functions, then the function $F_{Y_{i1}, \ldots, Y_{ik}}$ given by (1) is a joint distribution with marginals $F_{Y_{i1}}, \ldots, F_{Y_{ik}}$. The
copula $C$ may not be unique for discrete margins, however the practical limitations for statistical purposes are little, cf. Genest and Nešlehová (2007).

Owing to Sklar’s theorem, parametric families of copulas represent a powerful tool to describe the joint relationship between dependent random variables. Selecting the appropriate dependence within an assumed parametric copula family reduces to the selection of copula parameters, which correspond, for example, to a specific measure of association for the modeled random variables. Assuming $Y_{i1}, \ldots, Y_{ik}$ are continuous random variables with associated copula $C(\cdot; \alpha)$, one measure of association proposed by Joe (1990) is given by

$$
\tau_k = \frac{1}{2^k - 1} \left\{ 2^k \int_{[0,1]^k} C(\cdot; \alpha)dC(\cdot; \alpha) - 1 \right\}. \quad (2)
$$

Equation (2) is a generalized version of Kendall’s $\tau$, and hence establishes a correspondence between a scalar copula parameter $\alpha$ and the degree of dependence. More details and properties of this quantity, and another more traditional measure of concordance, can be found in Genest et al. (2011).

### 2.2 Design of experiments for copula models

In common with most work on optimal design of experiments, we base our criterion on the Fisher information matrix (FIM), the inverse of which provides an asymptotic approximation to the variance-covariance matrix of the maximum likelihood estimators of the model parameters.

Let $\zeta_i = (x_{i1}, \ldots, x_{ik}) \in \mathcal{X}^k$ denote the $k$ treatment vectors assigned to the units in block $i$ ($i = 1, \ldots, b$; $j = 1, \ldots, k$). We will work within a class of normalized block designs defined as

$$
\xi = \left\{ \frac{\zeta_1}{w_1}, \ldots, \frac{\zeta_n}{w_n} \right\}, \quad 0 < w_i \leq 1, \quad \sum_{i=1}^{n} w_i = 1,
$$

with $n \leq b$ distinct (support) blocks. As defined, $bw_i$ must be integer and represents the replication of the $i$th support block ($i = 1, \ldots, n$). Without loss of generality, we assume the first $n$ blocks in the design correspond to $\zeta_1, \ldots, \zeta_n$, with the remaining $b - n$ blocks being replicates.

We relax the assumption that $bw_i$ is integer to find so-called approximate or continuous designs; see also Cheng (1995) and Waite and Woods (2015). Let $\Xi$ denote the space of all possible designs of this form.

Denote the vector of responses from the $i$th block as $Y_i = (Y_{i1}, \ldots, Y_{ik})^T$, $i = 1, \ldots, b$, with corresponding expectation vector

$$
\eta_i = [\eta(x_{i1}; \beta), \ldots, \eta(x_{ik}; \beta)]^T,
$$

where $\eta(\cdot; \cdot)$ is a known function and $\beta = (\beta_1, \ldots, \beta_r)^T$ is a vector of unknown parameters requiring estimation. Denote the marginal distribution function for the $j$th entry in the block as $F_{Y_{ij}}(y_{ij}; x_{ij}, \beta)$, $j = 1, \ldots, k$, and denote the joint distribution, derived via a copula transformation, for the $k$ responses in the $i$th block as $C(F_{Y_{i1}}, \ldots, F_{Y_{ik}}; \alpha)$ where $\alpha = (\alpha_1, \ldots, \alpha_l)^T$ are unknown (copula) parameters.

The FIM $M(\zeta_i; \gamma)$ for the $i$th block is an $(r + l) \times (r + l)$ matrix with $vw$th element

$$
M(\zeta_i; \gamma)_{vw} = E \left[ -\frac{\partial^2}{\partial \gamma_v \partial \gamma_w} \log c_{Y_i}(\eta_i, \alpha) \right], \quad (3)
$$
where $\gamma = (\gamma_1, \ldots, \gamma_{r+l})^T = (\beta_1, \ldots, \beta_r, \alpha_1, \ldots, \alpha_l)^T$ and

$$c_{\gamma_i}(\eta_i, \alpha) = \frac{\partial^k}{\partial y_{i1} \ldots \partial y_{ik}} C (F_{Y_{i1}}, \ldots, F_{Y_{ik}}; \alpha)$$

is the joint density function represented through a copula $C$ in accordance with Equation [1]. The FIM for an approximate block design $\xi$ is then given by

$$M(\xi; \gamma) = \sum_{i=1}^{n} w_i M(\zeta_i; \gamma).$$

An optimal design $\xi^*$ maximizes a scalar function $\psi \{M(\xi; \gamma)\}$ of the information matrix. Previous work on optimal designs for copulas has focussed on finding completely randomized locally-optimal designs for multivariate responses, which can be considered as a block design where every unit within a block must receive the same treatment. Denman et al. (2011) found $D_{A}$-optimal designs for a bivariate response ($k = 2$) that maximized $\psi^D \{M(\xi; \gamma)\} = \det M(\xi; \gamma)$, and Perrone and Müller (2016) developed a corresponding equivalence theorem. These methods were extended to the local $D_{A}$-criterion, and, as a special case, for the $D_{s}$-criterion in Perrone et al. (2016). Other relevant uses of design of experiments in copula models are Deldossi et al. (2018) and Durante and Perrone (2016), but until now all relied on the availability of a single “best guess” vector of parameter values.

To overcome this dependence on assumed parameter values, here we adopt a pseudo-Bayesian approach to constructing block designs. Furthermore, our primary interest is typically in an $s$ meaningful linear combination of the parameters. Such combinations can be defined as elements of the variance-covariance matrix of the maximum likelihood estimator of $\gamma$. Most often the main interest is in an $s < (r+l)$-dimensional subset of the parameters. In such a case, a robust $D_s$-optimal block design can be found by maximizing

$$\Psi^D(\xi; G) = \int_{\Gamma} \log \det \{ M(\xi; \gamma) - M_{11} M_{22}^{-1} M_{12} \} \, dG(\gamma),$$

following the partition of the information matrix as

$$M(\xi; \gamma) = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.$$

Here, $M_{11}$ is the $(s \times s)$ partition related to the parameters of interest. This criterion follows as a special case of the $D_{A}$-criterion with $A^T = (I_s \ 0_{s \times (r+l-s)})$, with $I_s$ the $s \times s$ identity matrix and $0_{s \times (r+l-s)}$ the $s \times (r+l-s)$ zero matrix.

We evaluate a design $\xi$ via its Bayesian efficiencies under a given criterion, relative to an appropriate reference design $\xi^*$ (see, for example, Waite 2018). Under robust $D_s$-optimality, this efficiency is given by:

$$\text{eff}(\xi, \xi^*) = \left( \frac{\exp \int_B \log \det [M_{11}(\xi, \gamma) - M_{12}(\xi, \gamma) M_{22}^{-1}(\xi, \gamma) M_{12}(\xi, \gamma)] \, dF(\gamma)}{\exp \int_B \log \det [M_{11}(\xi^*, \gamma) - M_{12}(\xi^*, \gamma) M_{22}^{-1}(\xi^*, \gamma) M_{12}(\xi^*, \gamma)] \, dF(\gamma)} \right)^{1/s}.$$
We find designs that maximize (4) and (5) numerically using a version of the Fedorov-Wynn
algorithm (Wynn 1970; Fedorov 1971), as implemented in R package docopulae (Rappold 2018).

The optimality of a block design \( \xi^* \) under the robust \( D_A \)-criterion, regardless of how it
was found, can be assessed via application of the following Kiefer-Wolfowitz-type equivalence
theorem. The proof is similar to that for completely randomized experiments with multivariate
response, see Perrone et al. (2016) for the locally-optimal design case.

**Theorem 2.2.** The following properties are equivalent:

1. \( \xi^* \) is \( D_A \)-optimal;
2. for every \( \zeta \in \mathcal{X}^k \),
   \[
   \int B \text{tr} \left[ M(\xi^*; \gamma)^{-1} A (A^T M(\xi^*; \gamma)^{-1} A)^{-1} A^T M(\xi^*; \gamma)^{-1} M(\zeta; \gamma) \right] dG(\gamma) \leq s;
   \]
3. over all \( \xi \in \Xi \), the design \( \xi^* \) minimizes the function
   \[
   \max_{\zeta \in \mathcal{X}^k} \int B \text{tr} \left[ M(\xi^*; \gamma)^{-1} A (A^T M(\xi^*; \gamma)^{-1} A)^{-1} A^T M(\xi^*; \gamma)^{-1} M(\zeta; \gamma) \right] dG(\gamma),
   \]
where \( \Xi \) is the set of all possible block designs.

**2.3 Comparative example**

We demonstrate robust optimal block designs for copula models using a simple example from
Woods and van de Ven (2011), which allows comparison to the designs found by those authors
for a GEE model. We find robust designs for a single-factor log-linear regression model assuming
Poisson marginal distributions and quadratic linear predictor, implying \( \log \left\{ \eta(x; \beta) \right\} = \beta_0 + \beta_1 x + \beta_2 x^2 \). The prior distribution \( G \) is uniform on the parameter space \([-1, 1] \times [4, 5] \times [0.5, 1.5] \).

In line with our motivating example, we assume blocks of size \( k = 2 \) and intra-block dependence
defined according to one of the following bivariate copula functions.

1. **Product Copula**, which represents the independence case,
   \[
   C(u_1, u_2) = u_1 u_2,
   \]
   with generalized Kendall’s \( \tau \) of \( \tau_2 = 0 \).

2. **Clayton Copula**,\n   \[
   C_{\alpha}(u_1, u_2; \alpha) = \left[ \max \left( u_1^{-\alpha} + u_2^{-\alpha} - 1, 0 \right) \right]^{-\frac{1}{\alpha}},
   \]
   with \( \alpha \in (0, +\infty) \) and generalized \( \tau_2 = -\frac{\alpha}{\alpha+2} \).

3. **Gumbel Copula**,\n   \[
   C_{\alpha}(u_1, u_2; \alpha) = \exp \left( - \left[ (-\ln u_1)^\alpha + (-\ln u_2)^\alpha \right]^{\frac{1}{\alpha}} \right),
   \]
   with \( \alpha \in [1, +\infty) \) and generalized \( \tau_2 = \frac{\alpha-1}{\alpha} \).
The first copula is chosen for reference purposes; the latter two represent opposing dependencies in the tails (lower tail dependence for the Clayton versus upper tail dependence for the Gumbel). To isolate the effect of the copula structure from the strength of the dependence, we set $\alpha$ for each copula such that the values for Kendall’s $\tau$ coincide at three levels: $\tau_2 = \epsilon > 0, 1/10, 1/3$ respectively. Here $\epsilon = 10^{-9}$ is a small number to approximate the zero case, but avoid singularity issues.

To find robust $D$-optimal designs, objective function (4) was evaluated using quadrature (Gotwalt et al., 2009). Optimal designs under the Clayton and Gumbel copulas are shown in Figure 2 and demonstrate that increasing the generalized dependence (i.e. increasing $\tau_2$) leads to designs placing more weight on support blocks with points on the edge of the design space. All the designs display a “mirror-image” structure, with all design points having $x > 0$. These features are common in designs for Poisson regression (see Russell et al., 2009). The designs found under the Gumbel copula tend to include more support blocks but the pattern in the changes to these blocks as $\tau_2$ is increased is similar for both copulas.

Figure 2: Optimal designs for the comparative example; rows: Clayton and Gumbel copula; columns levels $\tau_2 = \epsilon > 0, 1/10, 1/3$.

For reference purposes the optimal design using the independence copula, i.e. an optimal design assuming no block effect, was evaluated. It showed little difference to setting the nominal level for $\tau_2 = 0$ for a particular copula. In particular the D-efficiencies for the Clayton and Gumbel model were 96.3% and 99.7% respectively. This efficiency expectedly decreases as the association within the block increases, for $\tau_2 = 1/3$ for instance it is already down to 65.0% and 61.3% respectively.

In Woods and van de Ven (2011), robust $D$-optimal designs were found under the same Poisson marginal models and prior distribution but with the dependence described using a GEE approach with an exchangeable correlation matrix and pairwise working correlation of
0.5. The optimal design found was given by:

\[ \xi^* = \left\{ \begin{array}{ccc}
(0.03, 1) & (1, 0.60) & (-0.40, 0.78) \\
0.355 & 0.310 & 0.335 
\end{array} \right\}. \] (6)

That is, for example, the first support block is \( \zeta_1 = (0.03, 1) \). This design is somewhat different in structure to the copula designs, without the same mirror structure. Quantitatively, the comparison shows the efficiencies under various scenarios given in Table 1. Surprisingly the design from Woods and van de Ven (2011) seems to be most compatible with an independence assumption.

| Independence | Clayton, \( \tau_2 = \epsilon > 0 \) | Clayton, \( \tau_2 = 1/3 \) | Gumbel, \( \tau_2 = \epsilon > 0 \) | Gumbel, \( \tau_2 = 1/3 \) |
|--------------|-------------------------------|-----------------|------------------|-----------------|
| 96.48%       | 89.85%                        | 84.41%          | 95.55%           | 92.96%          |

Table 1: D-efficiencies for design (6) from the GEE-approach designs under various copula models.

3 Application to the materials example

In this section we return to the materials testing example to find and assess designs for comparing six materials in block of size two under a variety of modelling assumptions. The measured response is binary, with each material sample either passing or failing a visual check. We label the five novel materials as “treatments”, with the reference material considered as a control. Marginally, we assume a logistic regression to model the differences between materials set up as

\[ Y_{ij} \sim \text{Bernoulli} \{ \eta(x_{ij}; \beta) \}; \quad \eta(x_{ij}; \beta) = \expit \left( \beta_0 + \sum_{l=1}^{5} \beta_l x_{ijl} \right), \]

where \( \expit(u) = 1/(1 + \exp(-u)) \), \( Y_{ij} \) is the binary response from the \( i \)th unit in the \( j \)th block \( (i = 1, 2; j = 1, \ldots, b) \), \( \eta(x_{ij}; \beta) \) is the associated probability of success, \( x_{ijl} \) is an indicator variable taking the value 1 if the \( i \)th unit in the \( j \)th block was assigned treatment \( l \) \((l = 1, \ldots, 5) \) and 0 otherwise, and \( \beta_0, \ldots, \beta_5 \) are unknown parameters to be estimated. Here, \( \beta_0 \) is the logit for the reference material, with \( \beta_l \) being the difference in expected response, on the logit scale, between the reference material and the \( l \)th novel material or treatment.

The choice of copula and the strength of intra-block association makes little difference to the design selected. However, assuming different marginal models and adopting a local or pseudo-Bayesian approach has a strong impact on the designs. Example designs for the Gumbel copula are shown in Figure 3.

With a null marginal model, i.e. \( \beta^T = (0, 0, 0, 0, 0, 0) \), when the response variance is constant, the locally D-optimal design contains all material combinations, excluding those blocks containing replicates of a single treatment. This design would also be optimal under a linear model with constant error variance. For different assumed parameter vectors, for example \( \beta^T = (0, -1, 2, -3, 4, -5) \), the optimal design contains only a few distinct treatment and treatment control combinations, with differing weights; here \((1,2),(3,4),(4,5)\) and \((5,6)\) are selected. The (pseudo)-Bayesian approach, assuming a continuous uniform prior on \([-1, 1] \times [-2, 0] \times [1, 3] \times [-4, -2] \times [3, 5] \times [-6, -4], so centred on \( \beta^T = (0, -1, 2, -3, 4, -5) \), adjusts the weighting of the support blocks to give more emphasis on comparing treatments 2 and 4 and 3 and 5. These pairs of treatments have differences to the control with the same sign.
Figure 3: Optimal designs for the materials testing example assuming a Gumbel copula with \( \tau_2 = 0.33 \); rows - local and pseudo-Bayesian; columns - assumed parameters or prior mean of \( \beta^T = (0, 0, 0, 0, 0, 0) \) and \( \beta^T = (0, -1, 2, -3, 4, -5) \), respectively.

4 Discussion

The modeling of block effects by copulas seems a natural choice and allows for elegant separation of the block and the marginal effects. Experimental designs for such models are now readily calculable. The pseudo-Bayesian \( D_A \)-optimality criterion was added to the \( \text{R} \) package \texttt{docopulae} version 0.4 (see Rappold, 2018) with the functions \texttt{wDsensitivity} and \texttt{wDeficiency}, both relying on a prespecified quadrature scheme for evaluation of the integrals. In this paper we have concentrated on finding designs to estimate the complete parameter vector but the implementation provides flexibility for checking for symmetry, model discrimination, etc., as investigated in Perrone et al. (2016).

Our examples are confined to the case \( k = 2 \). Whilst there is no theoretical necessity for that it is difficult to specify high-dimensional parametric copulas with a sufficient range of dependence, for details see the excellent survey of Nikoloulopoulos (2013). However, work on this issue would go well beyond the scope of this paper. It might also be interesting to contrast our findings with some known analytic results for blocks of size two as, for example, given in Cheng (1995) where a Gaussian copula is implicitly assumed.
Acknowledgements

We are grateful to Keith Warburton and Rob Ashmore from the UK Defence Science and Technology Laboratory for providing details of the materials testing example. W.G. Müller would like to acknowledge the hospitality of the Southampton Statistical Sciences Research Institute during his sabbatical, when this research was initiated. He was partially supported by project grants LIT-2017-4-SEE-001 funded by the Upper Austrian Government, and Austrian Science Fund (FWF): I 3903-N32 and D.C. Woods was partially supported by Fellowship EP/J018317/1 from the UK Engineering and Physical Sciences Research Council.

References

Atkinson, A. C., Donev, A. N., and Tobias, R. D. (2007), *Optimum Experimental Design, with SAS*, Oxford: Oxford University Press, 2nd ed.

Atkinson, A. C. and Ucinski, D. (2004), “Experimental design for time-dependent models with correlated observations,” *Studies in Nonlinear Dynamics and Econometrics*, 8, Article 13.

Atkinson, A. C. and Woods, D. C. (2015), “Designs for generalized linear models,” in *Handbook of Design and Analysis of Experiments*, eds. Dean, A. M., Morris, M. D., Stufken, J., and Bingham, D. R., Boca Raton: Chapman & Hall/CRC.

Bailey, R. A. (2007), “Designs for two-colour microarray experiments,” *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 56, 365–394.

Breslow, N. E. and Clayton, D. G. (1993), “Approximate inference in generalized linear mixed models,” *J. Am. Statist. Assoc.*, 88, 9–25.

Cheng, C. S. (1995), “Optimal regression designs under random block-effects models,” *Statist. Sinica*, 5, 485–497.

David, O. and Kempton, R. A. (1996), “Designs for Interference,” *Biometrics*, 52, 597–606.

Deldossi, L., Osmetti, S., and Tommasi, C. (2018), “Optimal Design to Discriminate between Rival Copula Models for a Bivariate Binary Response,” *TEST*.

Denman, N., McGree, J., Eccleston, J., and Duffull, S. (2011), “Design of experiments for bivariate binary responses modelled by Copula functions,” *Computational Statistics & Data Analysis*, 55, 1509–1520.

Dror, H. A. and Steinberg, D. M. (2006), “Robust experimental design for multivariate generalized linear models,” *Technometrics*, 48, 520–529.

— (2008), “Sequential experimental designs for generalized linear models,” *Journal of the American Statistical Association*, 103, 288–298.

Durante, F. and Perrone, E. (2016), “Asymmetric Copulas and Their Application in Design of Experiments,” in *On Logical, Algebraic, and Probabilistic Aspects of Fuzzy Set Theory*, eds. Saminger-Platz, S. and Mesiar, R., Springer International Publishing, vol. 336, pp. 157–172.

Fedorov, V. V. (1971), “The design of experiments in the multiresponse case,” *Theory of Probability and its Applications*, 16, 323–332.
Genest, C. and Nešlehová, J. (2007), “A Primer on Copulas for Count Data,” ASTIN Bulletin: The Journal of the International Actuarial Association, 37, 475–515.

Genest, C., Nešlehová, J., and Ben Ghorbal, N. (2011), “Estimators based on Kendall’s tau in multivariate copula models,” Australian & New Zealand Journal of Statistics, 53, 157–177.

Godolphin, J. (2018), “Designs with blocks of size two and application to microarray experiments,” The Annals of Statistics, 46, 2775–2805.

Gotwalt, C. M., Jones, B. A., and Steinberg, D. M. (2009), “Fast computation of designs robust to parameter uncertainty for nonlinear settings,” Technometrics, 51, 88–95.

Hughes-Oliver, J. M. (1998), “Optimal designs for nonlinear models with correlated errors,” in New Developments and Applications in Experimental Design, Institute of Mathematical Sciences, vol. 34 of IMS Lecture Notes - Monograph Series, pp. 163–174.

Joe, H. (1990), “Multivariate concordance,” Journal of Multivariate Analysis, 35, 12–30.

Kerr, K. (2012), “Optimality Criteria for the Design of 2-Color Microarray Studies,” Statistical applications in genetics and molecular biology, 11, Article–10.

King, J. and Wong, W. K. (2000), “Minimax D-optimal designs for the logistic model,” Biometrics, 56, 1263–1267.

Liang, K. Y. and Zeger, S. L. (1986), “Longitudinal data analysis using generalized linear models,” Biometrika, 73, 13–22.

Nelsen, R. B. (2007), An Introduction to Copulas (Springer Series in Statistics), Springer, 2nd ed.

Niaparast, M. and Schwabe, R. (2013), “Optimal design for quasi-likelihood estimation in Poisson regression with random coefficients,” Journal of Statistical Planning and Inference, 143, 296–306.

Nikoloulopoulos, A. (2013), “Copula-Based Models for Multivariate Discrete Response Data,” in Copulae in Mathematical and Quantitative Finance, eds. Jaworski, P., Durante, F., and Härdle, W. K., Springer Berlin Heidelberg, vol. 213 of Lecture Notes in Statistics, pp. 231–249.

Overstall, A. M. and Woods, D. C. (2017), “Bayesian design of experiments using approximate coordinate exchange,” Technometrics, 59, 458–470.

Perrone, E. and Müller, W. G. (2016), “Optimal designs for copula models,” Statistics, 50, 917–929.

Perrone, E., Rappold, A., and Müller, W. (2016), “D_s-optimality in copula models,” Statistical Methods & Applications, 26, 403–418.

Rappold, A. (2018), docopulae: Optimal designs for copula models, R package version 0.3.5.

Russell, K. G., Woods, D. C., Lewis, S. M., and Eccleston, J. A. (2009), “D-optimal designs for Poisson regression models,” Statistica Sinica, 19, 721–730.

Sklar, A. (1959), “Fonctions de répartition à n dimensions et leurs marges,” Publications de l’Institut de Statistique de Paris, 8, 229–231.
van de Ven, P. M. and Woods, D. C. (2014), “Optimal blocked minimum-support designs for non-linear models,” *Journal of Statistical Planning and Inference*, 144, 152–159.

Waite, T. W. (2018), “Singular prior distributions and ill-conditioning in Bayesian $D$-optimal design for several nonlinear models,” *Statistica Sinica*, 28, 505–525.

Waite, T. W. and Woods, D. C. (2015), “Designs for generalized linear models with random block effects via information matrix approximations,” *Biometrika*, 102, 677–693.

Woods, D. C., Lewis, S. M., Eccleston, J. A., and Russell, K. G. (2006), “Designs for generalized linear models with several variables and model uncertainty,” *Technometrics*, 48, 284–292.

Woods, D. C. and van de Ven, P. (2011), “Blocked Designs for Experiments With Correlated Non-Normal Response,” *Technometrics*, 53, 173–182.

Wynn, H. P. (1970), “The sequential generation of $D$-optimal experimental designs,” *The Annals of Mathematical Statistics*, 41, 1655–1664.