Supplementary Materials to "Model-Based Sampling Design for Multivariate Geostatistics"

Jie Li* and Dale L. Zimmerman†

Abstract

These supplementary materials provide supplemental tables and figures referenced in the paper, a description of the simulated annealing algorithm used to search for optimal designs in the large examples, and proofs of Theorems 1 and 2.

*Jie Li (Email: jieli@vt.edu) is Research Assistant Professor, Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24601.

†Dale L. Zimmerman (Email: dale-zimmerman@uiowa.edu) is Robert V. Hogg Professor, Department of Statistics and Actuarial Science, 241 Schaeffer Hall, University of Iowa, Iowa City, IA 52242.

1 Supplemental Tables

Supplemental Table 1: Relative MPEV-efficiencies of locally MPEV-optimal designs ($[\gamma_{MPEV}(D^*_{MPEV}(\boldsymbol{\theta});\boldsymbol{\theta})/\gamma_{MPEV}(D^*_{MPEV}(\boldsymbol{\theta}');\boldsymbol{\theta})] \times 100\%$, with $\boldsymbol{\theta}$ corresponding to rows and $\boldsymbol{\theta}'$ to columns) and the locally worst design ($[\gamma_{MPEV}(D^*_{MPEV}(\boldsymbol{\theta});\boldsymbol{\theta})/\max_D \gamma_{MPEV}(D;\boldsymbol{\theta})] \times 100\%$), for nine cases of cross-correlation and spatial correlation parameters in the Mat(0.5) examples.

		Toy example									
	ρ_c	0.2	0.2	0.2	0.5	0.5	0.5	0.8	0.8	0.8	locally worst
ρ_c	ρ	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8	design
0.2	0.2	100	98	98	99	100	98	99	99	95	65
0.2	0.5	99	100	100	89	97	100	89	89	93	28
0.2	0.8	92	100	100	75	89	100	75	75	85	10
0.5	0.2	93	91	91	100	99	91	100	100	95	60
0.5	0.5	99	99	99	95	100	99	95	95	97	30
0.5	0.8	92	100	100	81	91	100	81	81	90	11
0.8	0.2	66	64	64	100	92	64	100	100	91	38
0.8	0.5	77	75	75	100	91	75	100	100	96	20
0.8	0.8	85	88	88	98	87	88	98	98	100	11
		Large example									
	ρ_c	0.2	0.2	0.2	0.5	0.5	0.5	0.8	0.8	0.8	locally worst
ρ_c	ρ	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8	design
0.2	0.2	100	100	100	100	100	100	100	100	100	90
0.2	0.5	96	100	99	96	99	99	97	98	98	67
0.2	0.8	63	72	100	60	78	99	62	78	87	19
0.5	0.2	100	97	96	100	100	97	100	100	100	85
0.5	0.5	98	99	98	98	100	99	99	100	99	58
0.5	0.8	66	73	99	62	81	100	64	81	91	16
0.8	0.2	100	87	82	100	99	87	100	100	99	68
0.8	0.5	99	90	85	98	99	89	99	100	99	47
0.8	0.8	73	75	94	69	90	96	71	88	100	11

Supplemental Table 2: Relative CPE-efficiencies of locally CPE-optimal designs ($[\gamma_{CPE}(D^*_{CPE}(\boldsymbol{\theta});\boldsymbol{\theta})/\gamma_{CPE}(D^*_{CPE}(\boldsymbol{\theta}');\boldsymbol{\theta})] \times 100\%$, with $\boldsymbol{\theta}$ corresponding to rows and $\boldsymbol{\theta}'$ to columns) and the locally worst design ($[\gamma_{CPE}(D^*_{CPE}(\boldsymbol{\theta});\boldsymbol{\theta})/\max_D \gamma_{CPE}(D;\boldsymbol{\theta})] \times 100\%$), for nine cases of cross-correlation and spatial correlation parameters in the Mat(0.5) examples.

		Toy example									
	ρ_c	0.2	0.2	0.2	0.5	0.5	0.5	0.8	0.8	0.8	locally worst
ρ_c	ρ	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8	design
0.2	0.2	100	98	44	100	98	44	100	98	44	< 1
0.2	0.5	99	100	81	99	100	81	99	100	81	< 1
0.2	0.8	90	97	100	90	97	100	90	97	100	< 1
0.5	0.2	100	98	44	100	98	44	100	98	44	< 1
0.5	0.5	99	100	81	99	100	81	99	100	81	< 1
0.5	0.8	90	97	100	90	97	100	90	97	100	< 1
0.8	0.2	100	98	44	100	98	44	100	98	44	< 1
0.8	0.5	99	100	81	99	100	81	99	100	81	< 1
0.8	0.8	90	97	100	90	97	100	90	97	100	< 1
		Large example									
	ρ_c	0.2	0.2	0.2	0.5	0.5	0.5	0.8	0.8	0.8	locally worst
ρ_c	ρ	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8	design
0.2	0.2	100	91	47	100	91	46	100	87	37	< 1
0.2	0.5	79	100	66	76	100	66	84	100	61	< 1
0.2	0.8	44	79	100	23	78	100	51	80	97	< 1
0.5	0.2	98	90	46	100	90	45	99	86	36	< 1
0.5	0.5	78	99	66	75	100	65	83	99	60	< 1
0.5	0.8	44	79	100	23	78	100	51	79	97	< 1
0.8	0.2	100	91	46	100	91	46	100	87	37	< 1
0.8	0.5	79	100	66	76	100	66	84	100	61	< 1
0.8	0.8	45	81	100	24	80	100	53	82	100	< 1

Supplemental Table 3: Relative MEPEV-efficiencies of locally MEPEV-optimal designs ($[\gamma_{MEPEV}(D^*_{MEPEV}(\boldsymbol{\theta});\boldsymbol{\theta})/\gamma_{MEPEV}(D^*_{MEPEV}(\boldsymbol{\theta}');\boldsymbol{\theta})] \times 100\%$, with $\boldsymbol{\theta}$ corresponding to rows and $\boldsymbol{\theta}'$ to columns) and the locally worst design ($[\gamma_{MEPEV}(D^*_{MEPEV}(\boldsymbol{\theta});\boldsymbol{\theta})/\max_D \gamma_{MEPEV}(D;\boldsymbol{\theta})] \times 100\%$), for nine cases of cross-correlation and spatial correlation parameters in the Mat(0.5) examples.

		Toy example									
	ρ_c	0.2	0.2	0.2	0.5	0.5	0.5	0.8	0.8	0.8	locally worst
ρ_c	ρ	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8	design
0.2	0.2	100	53	< 1	100	53	< 1	100	53	< 1	< 1
0.2	0.5	70	100	59	70	100	59	70	100	59	3
0.2	0.8	21	65	100	21	65	100	21	65	100	< 1
0.5	0.2	100	50	< 1	100	50	< 1	100	50	< 1	< 1
0.5	0.5	72	100	61	72	100	61	72	100	61	2
0.5	0.8	21	66	100	21	66	100	21	66	100	< 1
0.8	0.2	100	46	< 1	100	46	< 1	100	46	< 1	< 1
0.8	0.5	77	100	65	77	100	65	77	100	65	1
0.8	0.8	21	65	100	21	65	100	21	65	100	< 1
		Large example									
	ρ_c	0.2	0.2	0.2	0.5	0.5	0.5	0.8	0.8	0.8	locally worst
ρ_c	ρ	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8	design
0.2	0.2	100	87	< 1	100	89	< 1	100	85	33	< 1
0.2	0.5	95	100	34	95	100	48	95	100	69	< 1
0.2	0.8	44	49	100	37	48	95	41	48	89	< 1
0.5	0.2	99	87	< 1	100	89	5	100	85	33	< 1
0.5	0.5	94	100	46	95	100	59	96	100	72	< 1
0.5	0.8	45	52	99	39	50	100	43	50	94	< 1
0.8	0.2	97	87	< 1	99	89	20	100	85	35	< 1
0.8	0.5	93	100	62	94	100	73	96	100	76	< 1
0.8	0.8	47	55	90	41	54	93	46	53	100	< 1

Supplemental Table 4: Relative MEPEV-efficiencies of (a) the locally MPEV-optimal designs (i.e. $[\gamma_{MEPEV}(D^*_{MEPEV}(\boldsymbol{\theta});\boldsymbol{\theta})/\gamma_{MEPEV}(D^*_{MPEV}(\boldsymbol{\theta});\boldsymbol{\theta})] \times 100\%$), and (b) the collocated maximin LHD, for nine cases of cross-correlation and spatial correlation parameters in the Mat(0.5) examples.

ρ_c	ρ	Toy example (a)	Toy example (b)	Large example (a)	Large example (b)
0.2	0.2	14	7	< 1	< 1
0.2	0.5	55	47	38	17
0.2	0.8	95	50	91	50
0.5	0.2	5	13	< 1	< 1
0.5	0.5	54	50	25	17
0.5	0.8	73	50	94	53
0.8	0.2	5	27	< 1	< 1
0.8	0.5	50	59	13	17
0.8	0.8	63	50	101	56

Supplemental Table 5: Relative MPEV-efficiencies of locally MPEV-optimal collocated designs (i.e., $[\gamma_{MPEV}(D^*_{MPEV}(\boldsymbol{\theta});\boldsymbol{\theta})/\gamma_{MPEV}(D^*_{MPEV_{collocated}};\boldsymbol{\theta})]\times 100\%$), for toy examples corresponding to seven spatial processes and nine cases of cross-correlation and spatial correlation parameters.

ρ_c	ρ	Mat(0.5)	Mat(1.5)	Mat(∞)	NS1	NS2	NS3	NS4
0.2	0.2	99	98	95	98	98	76	97
0.2	0.5	89	84	67	100	99	54	95
0.2	0.8	75	59	28	100	99	42	97
0.5	0.2	100	100	99	89	88	78	87
0.5	0.5	95	89	70	94	86	52	86
0.5	0.8	81	65	30	100	72	41	99
0.8	0.2	100	100	100	76	73	89	73
0.8	0.5	100	100	79	80	68	46	73
0.8	0.8	98	84	39	93	54	39	81

Supplemental Table 6: Relative MPEV-efficiencies of locally ostensibly MPEV-optimal collocated designs (i.e., $[\gamma_{MPEV}(D^*_{MPEV}(\boldsymbol{\theta});\boldsymbol{\theta})/\gamma_{MPEV}(D^*_{MPEV_{collocated}};\boldsymbol{\theta})]\times 100\%$), for large examples corresponding to seven spatial processes and nine cases of cross-correlation and spatial correlation parameters.

ρ_c	ρ	Mat(0.5)	Mat(1.5)	Mat(∞)	NS1	NS2	NS3	NS4
0.2	0.2	100	100	100	100	100	100	100
0.2	0.5	97	98	100	99	97	88	97
0.2	0.8	86	82	80	90	87	72	84
0.5	0.2	100	100	100	100	100	100	99
0.5	0.5	99	100	100	100	100	90	97
0.5	0.8	89	82	83	92	90	80	82
0.8	0.2	100	100	100	100	100	98	99
0.8	0.5	100	100	100	100	100	82	94
0.8	0.8	97	89	92	93	93	76	84

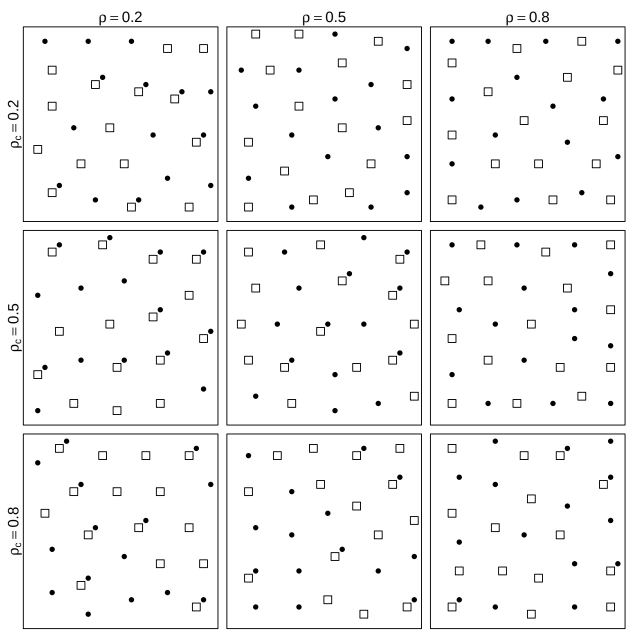
Supplemental Table 7: Relative MEPEV-efficiencies of locally ostensibly MEPEV-optimal collocated designs (i.e., $[\gamma_{MEPEV}(D^*_{MEPEV}(\boldsymbol{\theta});\boldsymbol{\theta})/\gamma_{MEPEV}(D^*_{MEPEV_{collocated}};\boldsymbol{\theta})]\times 100\%$), for large examples corresponding to seven spatial processes and nine cases of cross-correlation and spatial correlation parameters.

ρ_c	ρ	Mat(0.5)	Mat(1.5)	Mat(∞)	NS1	NS2	NS3	NS4
0.2	0.2	100	100	100	100	99	100	93
0.2	0.5	100	100	100	100	99	98	99
0.2	0.8	91	88	91	93	89	87	87
0.5	0.2	100	100	100	99	100	100	99
0.5	0.5	100	100	100	100	100	99	99
0.5	0.8	93	86	92	90	89	87	88
0.8	0.2	100	100	100	100	99	100	99
0.8	0.5	100	100	100	100	100	98	99
0.8	0.8	100	92	94	90	93	98	78

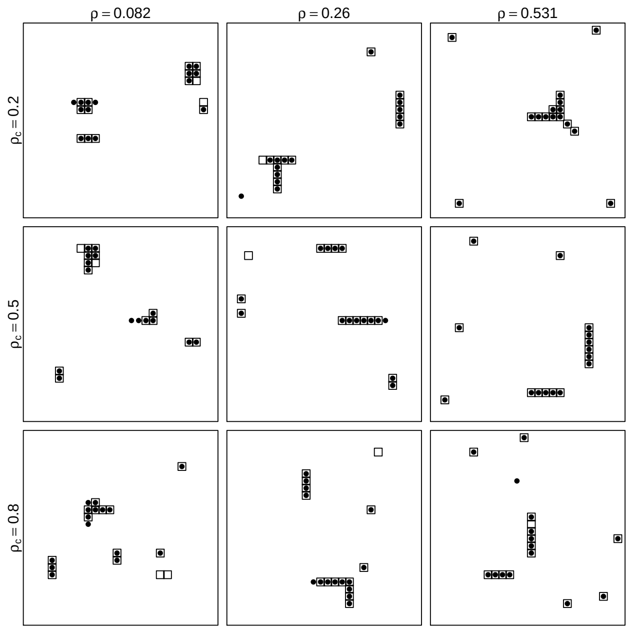
Supplemental Table 8: Optimality of balanced or unbalanced designs, for the NS3 toy example with $n_1+n_2 = 8$ and $3 \leq n_1 \leq 5$ and nine cases of cross-correlation and spatial correlation parameters. A table entry of $*$ indicates that the locally optimal design is unbalanced (in every such case $n_1 = 3$); otherwise, it is balanced.

ρ_c	ρ	MPEV	CPE	MEPEV
0.2	0.2			$*$
0.2	0.5	$*$		
0.2	0.8	$*$	$*$	
0.5	0.2			$*$
0.5	0.5	$*$		$*$
0.5	0.8	$*$	$*$	
0.8	0.2		$*$	$*$
0.8	0.5			$*$
0.8	0.8	$*$	$*$	

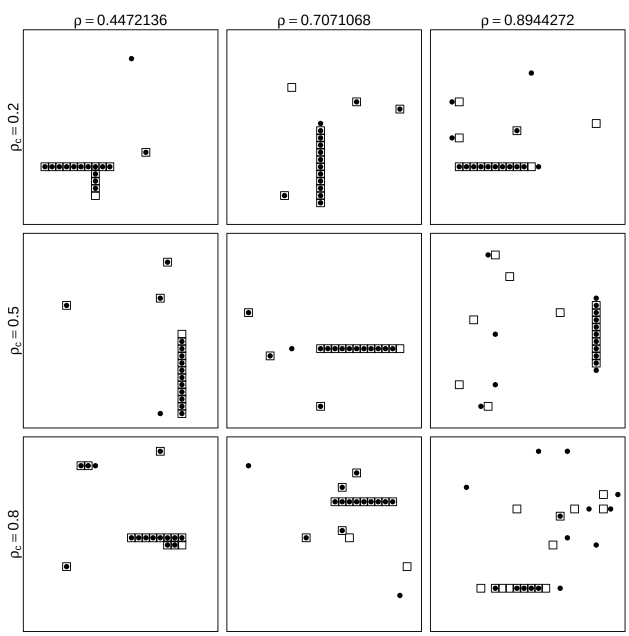
2 Supplemental Figures



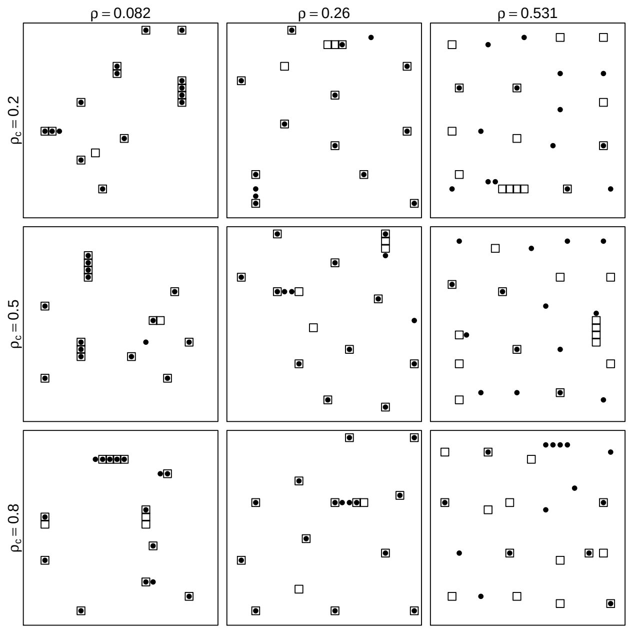
Supplemental Figure 1: Locally ostensibly MPEV-optimal designs for the NS4 large example.



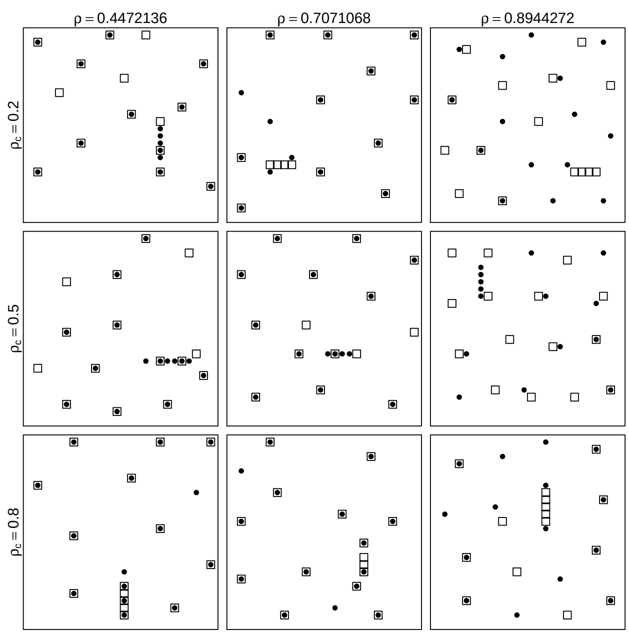
Supplemental Figure 2: Locally ostensibly CPE-optimal designs for the Mat(1.5) large example.



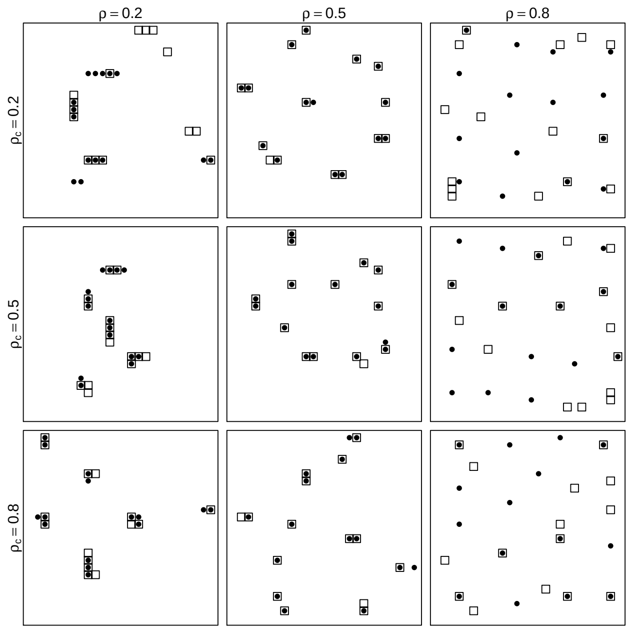
Supplemental Figure 3: Locally ostensibly CPE-optimal designs for the Mat(∞) large example.



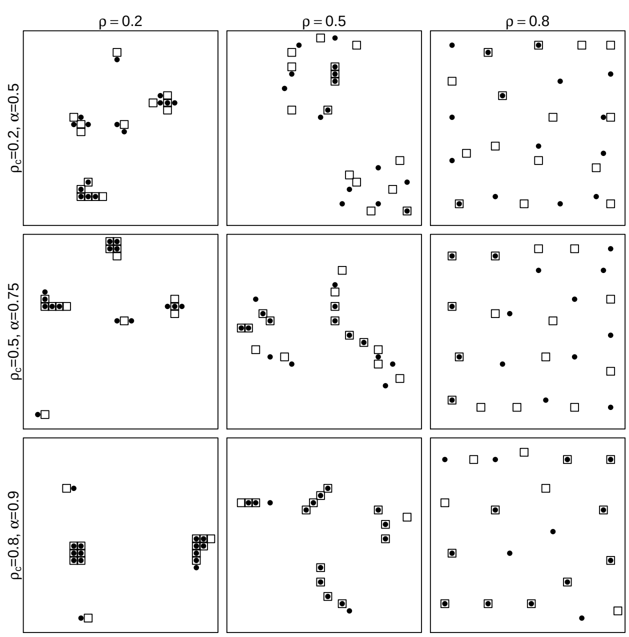
Supplemental Figure 4: Locally ostensibly MEPEV-optimal designs for the Mat(1.5) large example.



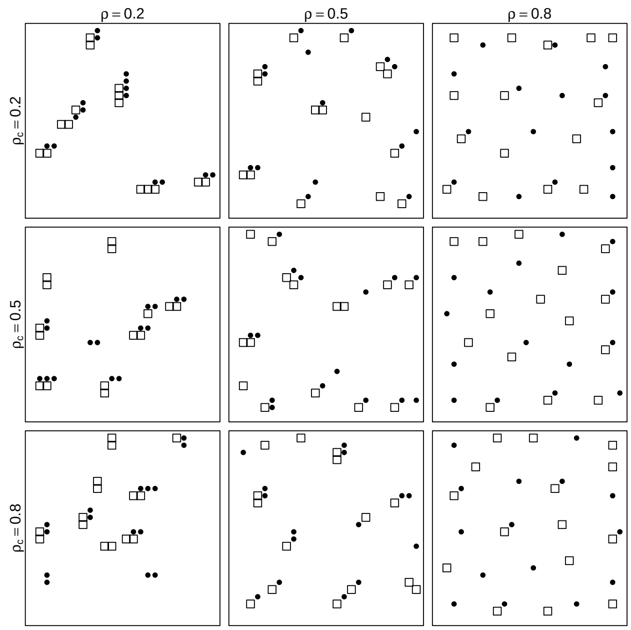
Supplemental Figure 5: Locally ostensibly MEPEV-optimal designs for the Mat(∞) large example.



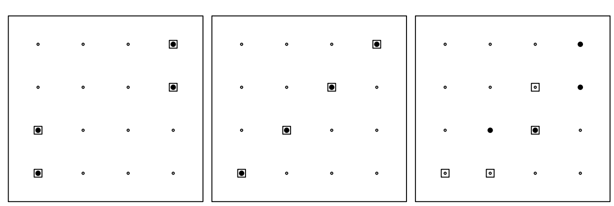
Supplemental Figure 6: Locally ostensibly MEPEV-optimal designs for the NS1 large example.



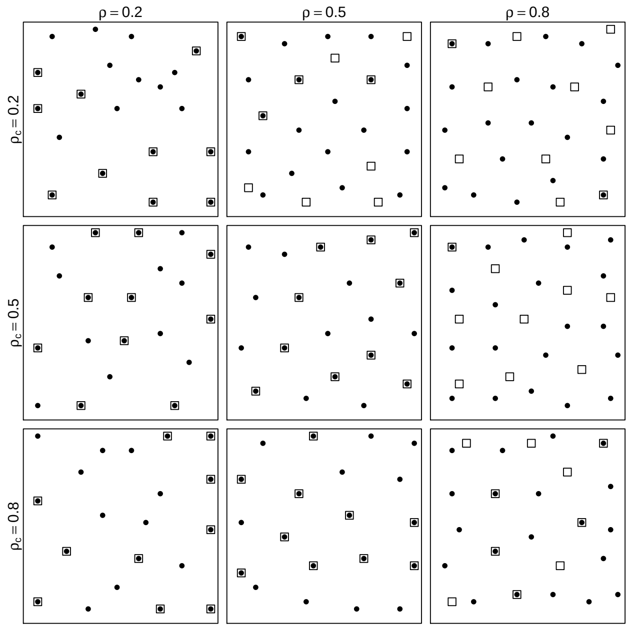
Supplemental Figure 7: Locally ostensibly MEPEV-optimal designs for the NS2 large example.



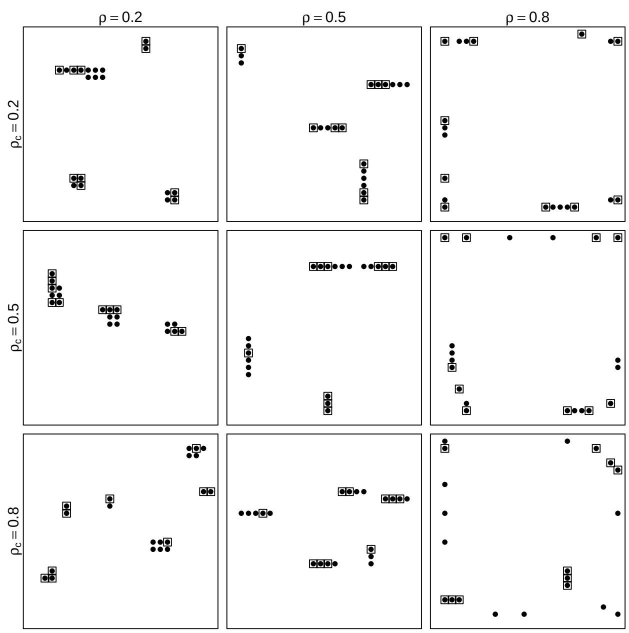
Supplemental Figure 8: Locally ostensibly MEPEV-optimal designs for the NS4 large example.



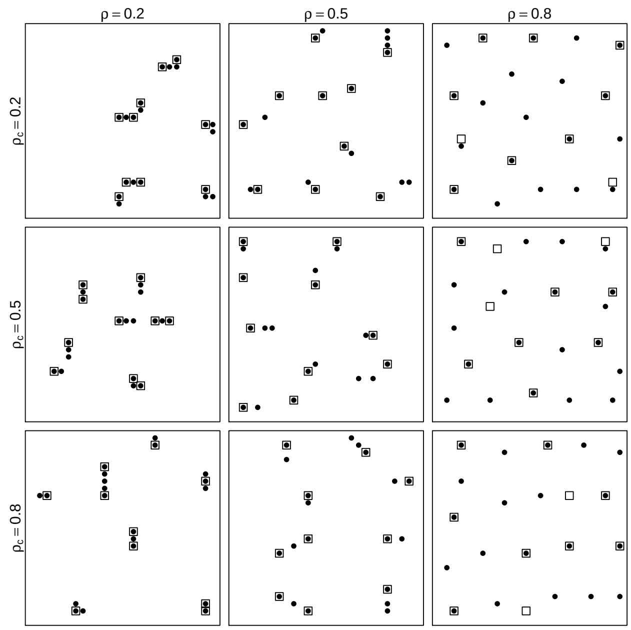
Supplemental Figure 9: The superposition of CPE-optimal univariate designs, the CPE-optimal collocated design, and the CPE-optimal design (left to right) for the NS4 toy example when $\rho_c = \rho = 0.2$.



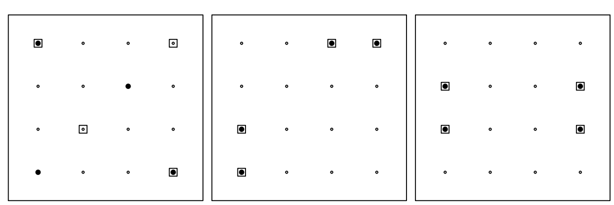
Supplemental Figure 10: Locally ostensibly MPEV-optimal unbalanced designs ($n_1 = 20, n_2 = 10$) for the NS2 large example.



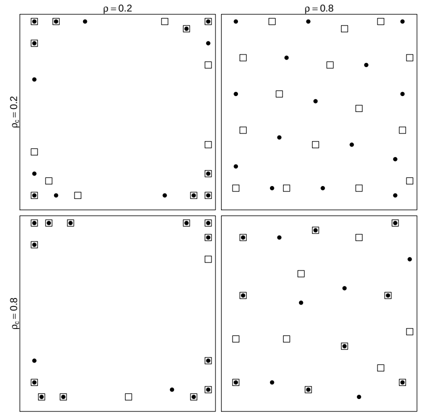
Supplemental Figure 11: Locally ostensibly CPE-optimal unbalanced designs ($n_1 = 20, n_2 = 10$) for the NS2 large example.



Supplemental Figure 12: Locally ostensibly MEPEV-optimal unbalanced designs ($n_1 = 20, n_2 = 10$) for the NS2 large example.



Supplemental Figure 13: Locally MPEV-, CPE-, and MEPEV-optimal compromise designs (left to right), using a desirability function, for the Mat(0.5) toy example.



Supplemental Figure 14: Locally ostensibly MPEV-optimal designs for the Mat(0.5) large example with planar mean.

3 Simulated Annealing Algorithm

For general use, we propose the following simulated annealing algorithm (SAA) to search for the locally optimal design for a fixed vector $\boldsymbol{\theta}$ of covariance parameters. The algorithm consists of the following steps:

1. Randomly choose a design D_0, then compute $\gamma(D_0;\boldsymbol{\theta})$, where $\gamma(\cdot)$ is the corresponding criterion function that we wish to minimize. Set the initial values for the "cooling factor" and "distance factor", τ_0 and h_{max_0}, respectively.

2. Randomly select a point in D_0 and move it h units in a randomly selected direction, where h has a uniform distribution on $[0, h_{max_0}]$. If the move takes the point to a location outside the design region, then the point is returned to its original location and a new move is generated at random until the moved point falls into the design region. Call this new design D_1.

3. Calculate $\gamma(D_1;\boldsymbol{\theta})$. Then the transition $D_0 \to D_1$ is accepted with the following probability:

$$P_\tau(D_0 \to D_1) = \begin{cases} 1, & \text{if } \gamma(D_1;\boldsymbol{\theta}) \leq \gamma(D_0;\boldsymbol{\theta}); \\ \exp\{[\gamma(D_0;\boldsymbol{\theta}) - \gamma(D_1;\boldsymbol{\theta})]/\tau_0\}, & \text{if } \gamma(D_1;\boldsymbol{\theta}) > \gamma(D_0;\boldsymbol{\theta}). \end{cases}$$

4. Repeat steps 2-3 for N_{sim} (sufficiently large) times.

5. At the end of the N_{sim} repetitions, reduce the "cooling factor" and "distance factor" to $\tau_{k+1} = \alpha_\tau \tau_k$, $h_{max_{k+1}} = \alpha_h h_{max_k}$.

6. Repeat steps 2-5 for N_t times.

In this algorithm, τ_0 was set so that 95% or more of the moves would be accepted before the first cooling step; α_τ was set so that the acceptance rate would be below 0.1% on the final step; h_{max_0} was set at half the length of the region to be sampled; and α_h was set such that the final h_{max} was 0.5 unit.

For our featured examples, the set of candidate design points is a finite rectangular grid so we can only move a point in four directions: north, south, east and west. Also, the distance to be moved is generated from a discrete uniform distribution, i.e., the distance must be an integer multiple of the unit spacing. In all cases, we chose $N_{sim} = 1000$ and $N_t = 1000$.

In our larger example, it took a few hours for the algorithm to converge to a (local) minimum, given a specific set of values for ρ_c and ρ. Different starting designs were used, including regular square grid, random, and clustered designs, each in completely collocated, completely disjoint, and mixed forms. Regardless of the design criterion, the starting design had no discernible effect on the qualitative characteristics and very little effect on either the criterion value of the final design or the number of iterations required to achieve convergence. Thus, the results are extremely robust to the choice of starting design.

As a check, we applied the SAA to the toy examples. In every case, the SAA converged to the same locally optimal design (or one within the same equivalence class) obtained by complete enumeration. However, in some cases it took longer to find this design using the SAA than by complete enumeration. A potential improvement of the SAA is to adaptively restrict the allowable directions a point may be moved, so as to eliminate the need to discard beyond-the-boundary moves from points on or near the boundary.

4 Proof of Theorem 1

Proof of (a):

Without loss of generality assume $m = 2$. Since the cross-correlation is zero,

$$\mathbf{\Sigma}_{00} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}, \qquad \mathbf{C}_0 = \begin{pmatrix} \sigma_1^2 \mathbf{c}_1 & \mathbf{0} \\ \mathbf{0} & \sigma_2^2 \mathbf{c}_2 \end{pmatrix}, \qquad \text{and} \quad \mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 \mathbf{R}_{11} & \mathbf{0} \\ \mathbf{0} & \sigma_2^2 \mathbf{R}_{22} \end{pmatrix}$$

for certain vectors $\mathbf{c}_1$ and $\mathbf{c}_2$ and correlation matrices $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$. Substitution into (2) then establishes that the off-diagonal elements of $\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D)$ are zero, so that $|\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D)|$ is the product of its main diagonal elements, which are nonnegative. Now recall the well-known fact that the best linear unbiased predictor of a variable at a site at which it is observed is merely the observed value of the variable at that site. Thus if $\mathbf{s}_0 \in D$, then at least one of the main diagonal elements of $\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D)$ is zero and thus $|\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D)| = 0$. For $\mathbf{s}_0 \notin D$, $|\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D)| > 0$. The maximum of $|\mathbf{V}(\mathbf{s}, \boldsymbol{\theta}, D)|$ over all $\mathbf{s} \in \mathcal{P}$ therefore is minimized by making $D \cap \mathcal{P}$ as large as possible, i.e. by imposing no collocation whatsoever. This completes the proof of part (a).

Proof of (b):

Again assume $m = 2$ without loss of generality. Consider first the value of $|\mathbf{V}|$ when the design is completely collocated, in which case $\mathbf{\Sigma} = \mathbf{\Sigma}_{00} \otimes \mathbf{I}_n$. If $\mathbf{s}_0 \in D$, then $\mathbf{C}_0 = \mathbf{\Sigma}_{00} \otimes \mathbf{e}_i$ for some $i = 1, \ldots, n$, where $\mathbf{e}_i$ is the $n \times 1$ vector with 1 as its ith element and zeros elsewhere, and

$$\begin{aligned} \mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D) &= \mathbf{\Sigma}_{00} - (\mathbf{\Sigma}_{00} \otimes \mathbf{e}_i)'(\mathbf{\Sigma}_{00} \otimes \mathbf{I}_n)^{-1}(\mathbf{\Sigma}_{00} \otimes \mathbf{e}_i) + [\mathbf{I}_2 - (\mathbf{I}_2 \otimes \mathbf{1}_n)'(\mathbf{\Sigma}_{00} \otimes \mathbf{I}_n)^{-1}(\mathbf{\Sigma}_{00} \otimes \mathbf{e}_i)]' \\ &\quad \times [(\mathbf{I}_2 \otimes \mathbf{1}_n)'(\mathbf{\Sigma}_{00} \otimes \mathbf{I}_n)^{-1}(\mathbf{I}_2 \otimes \mathbf{1}_n)]^{-1}[\mathbf{I}_2 - (\mathbf{I}_2 \otimes \mathbf{1}_n)'(\mathbf{\Sigma}_{00} \otimes \mathbf{I}_n)^{-1}(\mathbf{\Sigma}_{00} \otimes \mathbf{e}_i)] \\ &= \mathbf{0}. \end{aligned}$$

If, on the other hand, $\mathbf{s}_0 \notin D$, then $\mathbf{C}_0 = \mathbf{0}$ and

$$\begin{aligned} \mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D) &= \mathbf{\Sigma}_{00} + \mathbf{I}_2'[(\mathbf{I}_2 \otimes \mathbf{1}_n)'(\mathbf{\Sigma}_{00} \otimes \mathbf{I}_n)^{-1}(\mathbf{I}_2 \otimes \mathbf{1}_n)]^{-1}\mathbf{I}_2 \\ &= \left(\frac{n+1}{n}\right) \mathbf{\Sigma}_{00}. \end{aligned}$$

Thus when the design is completely collocated,

$$\max_{\mathbf{s} \in \mathcal{P}} |\mathbf{V}(\mathbf{s}, \boldsymbol{\theta}, D)| = \sigma_1^2 \sigma_2^2 (1 - \rho_c^2) \left(\frac{n+1}{n}\right)^2. \tag{1}$$

If the design is not completely collocated, then, letting $n - n_0$ denote the number of collocated sites in the design, the sites can be ordered in such a way that

$$\mathbf{\Sigma} = \begin{pmatrix} \mathrm{diag}(\sigma_1^2, \sigma_2^2) \otimes \mathbf{I}_{n_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_{00} \otimes \mathbf{I}_{n-n_0} \end{pmatrix} \quad \text{and} \quad \mathbf{X} = \begin{pmatrix} \mathbf{I}_2 \otimes \mathbf{1}_{n_0} \\ \mathbf{I}_2 \otimes \mathbf{1}_{n-n_0} \end{pmatrix}.$$

Furthermore, since $\mathrm{card}(\mathcal{P}) > 2n$ by assumption, there is at least one site $\mathbf{s}_0 \in \mathcal{P}$ for which

$\mathbf{s}_0 \notin D$. For such a site, $\mathbf{C}_0 = \mathbf{0}$ and thus $\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D) = \boldsymbol{\Sigma}_{00} + [n_0 \text{diag}(1/\sigma_1^2, 1/\sigma_2^2) + (n - n_0)\boldsymbol{\Sigma}_{00}^{-1}]^{-1}$. After some algebra, we obtain

$$|\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D)| = \sigma_1^2 \sigma_2^2 (n^2 - n_0^2 \rho_c^2)^{-2} [(n^2 + n - n_0^2 \rho_c^2 - n_0 \rho_c^2)^2 - \rho_c^2 (n^2 + n - n_0^2 \rho_c^2 - n_0)^2]. \quad (2)$$

After considerably more algebra and an appropriate factorization, it can be shown that the term in brackets in (2) is equal to $(1 - \rho_c^2)(n^2 - n_0^2 \rho_c^2)[(n+1)^2 - n_0^2 \rho_c^2]$. Thus

$$|\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D)| = \sigma_1^2 \sigma_2^2 (1 - \rho_c^2) \left(\frac{(n+1)^2 - n_0^2 \rho_c^2}{n^2 - n_0^2 \rho_c^2} \right),$$

which, for $\rho_c \neq 0$, is strictly larger than (1). This completes the proof of part (b).

5 Proof of Theorem 2

Proof of (a):

The distinct sites of a collocated design may be represented as $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n$. This, together with the specified conditions on the mean function, implies that $\mathbf{X}_1 = \mathbf{X}_2 = \cdots = \mathbf{X}_n$ and thus that $\mathbf{X} = \mathbf{I}_m \otimes \mathbf{X}_1$. Similarly, $\mathbf{X}_0 = \mathbf{I}_m \otimes [\mathbf{x}_1(\mathbf{s}_0)]'$. Furthermore, the separability of the process yields $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{00} \otimes \mathbf{R}$ and $\mathbf{C}_0 = \boldsymbol{\Sigma}_{00} \otimes \mathbf{r}_0$, where $\mathbf{R}$ is the $n \times n$ matrix of spatial correlations among the observations at $\mathbf{s}_1, \ldots, \mathbf{s}_n$ on any one of the m variables and $\mathbf{r}_0$ is the $n \times 1$ vector of spatial correlations between the observations at $\mathbf{s}_1, \ldots, \mathbf{s}_n$ on any variable and that same variable at $\mathbf{s}_0$. Then, using equation (2) of Section 3.1, the covariance matrix of prediction errors associated with the BLUP of $\mathbf{Z}(\mathbf{s}_0)$ is

$$\begin{aligned}
\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D) &= \boldsymbol{\Sigma}_{00} - (\boldsymbol{\Sigma}_{00} \otimes \mathbf{r}_0)'(\boldsymbol{\Sigma}_{00} \otimes \mathbf{R})^{-1}(\boldsymbol{\Sigma}_{00} \otimes \mathbf{r}_0) \\
&\quad + [(\mathbf{I}_m \otimes [\mathbf{x}_1(\mathbf{s}_0)]') - (\mathbf{I}_m \otimes \mathbf{X}_1)'(\boldsymbol{\Sigma}_{00} \otimes \mathbf{R})^{-1}(\boldsymbol{\Sigma}_{00} \otimes \mathbf{r}_0)]' \\
&\quad \times [(\mathbf{I}_m \otimes \mathbf{X}_1)'(\boldsymbol{\Sigma}_{00} \otimes \mathbf{R})^{-1}(\mathbf{I}_m \otimes \mathbf{X}_1)]^{-1} \\
&\quad \times [(\mathbf{I}_m \otimes [\mathbf{x}_1(\mathbf{s}_0)]') - (\mathbf{I}_m \otimes \mathbf{X}_1)'(\boldsymbol{\Sigma}_{00} \otimes \mathbf{R})^{-1}(\boldsymbol{\Sigma}_{00} \otimes \mathbf{r}_0)] \\
&= \boldsymbol{\Sigma}_{00} \otimes [1 - \mathbf{r}_0'\mathbf{R}^{-1}\mathbf{r}_0 + (\mathbf{x}_1(\mathbf{s}_0) - \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{r}_0)'(\mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1)^{-1}(\mathbf{x}_1(\mathbf{s}_0) - \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{r}_0)] \\
&= \boldsymbol{\Sigma}_{00} \otimes [v_1(\mathbf{s}_0, \boldsymbol{\theta}, D)/C_{11}(\mathbf{s}_0, \mathbf{s}_0; \boldsymbol{\theta})],
\end{aligned}$$

where $v_1(\mathbf{s}_0, \boldsymbol{\theta}, D)$ is the prediction error variance associated with the univariate BLUP of the first component of $\mathbf{Z}(\mathbf{s}_0)$. Thus $|\mathbf{V}(\mathbf{s}_0, \boldsymbol{\theta}, D)| = |\boldsymbol{\Sigma}_{00}| \cdot [v_1(\mathbf{s}_0, \boldsymbol{\theta}, D)/C_{11}(\mathbf{s}_0, \mathbf{s}_0; \boldsymbol{\theta})]^m$, so that the minimization of $\max_{\mathbf{s} \in \mathcal{P}} |\mathbf{V}(\mathbf{s}, \boldsymbol{\theta}, D)|$ over the design space is equivalent to the minimization of $\max_{\mathbf{s} \in \mathcal{P}} v_1(\mathbf{s}, \boldsymbol{\theta}, D)$. This completes the proof of part (a).

Proof of (b):

As noted in the proof of part (a), $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{00} \otimes \mathbf{R}$. Since $\boldsymbol{\Sigma}_{00}$ is symmetric, we have that $\boldsymbol{\Sigma}_{00} = \boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}'$, where $\boldsymbol{\Lambda}$ is a diagonal matrix whose positive elements $\lambda_1, \ldots, \lambda_m$ are the eigenvalues of $\boldsymbol{\Sigma}_{00}$ and $\boldsymbol{\Omega} = (\omega_{ij})$ is the corresponding matrix of orthonormal eigenvectors. Let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)'$, let $\boldsymbol{\omega}$ be the $m(m-1)/2 \times 1$ vector of below-diagonal elements of $\boldsymbol{\Omega}$, let $\boldsymbol{\rho}$ be the $s \times 1$ vector of spatial correlation parameters, and let $\boldsymbol{\xi} = (\xi_i) = (\boldsymbol{\lambda}', \boldsymbol{\omega}', \boldsymbol{\rho}')'$. For the REML information matrix $\mathbf{I}(\boldsymbol{\theta}, D)$, we have $\mathbf{I}(\boldsymbol{\theta}, D) = \mathbf{J}\mathbf{I}(\boldsymbol{\xi}, D)\mathbf{J}'$ where $\mathbf{J} = \partial\boldsymbol{\xi}/\partial\boldsymbol{\theta}$ is not affected by the design. Thus the CPE-optimal design is that which minimizes $1/|\mathbf{I}(\boldsymbol{\xi}, D)|$. The (i,j)th element of $\mathbf{I}(\boldsymbol{\xi}, D)$ is given by $\frac{1}{2}\mathrm{tr}(\mathbf{P}\frac{\partial\mathbf{\Sigma}}{\partial\xi_i}\mathbf{P}\frac{\partial\mathbf{\Sigma}}{\partial\xi_j})$, where $\mathbf{P} = (\mathbf{\Omega}\mathbf{\Lambda}^{-1}\mathbf{\Omega}') \otimes \mathbf{P}_u$, $\mathbf{P}_u = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{X}_1(\mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{R}^{-1}$, and

$$\frac{\partial\mathbf{\Sigma}}{\partial\xi_i} = \begin{cases} \mathbf{\Omega}\frac{\partial\mathbf{\Lambda}}{\partial\xi_i}\mathbf{\Omega}' \otimes \mathbf{R}, & i = 1,\ldots,m; \\ \left(\frac{\partial\mathbf{\Omega}}{\partial\xi_i}\mathbf{\Lambda}\mathbf{\Omega}' + \mathbf{\Omega}\mathbf{\Lambda}\frac{\partial\mathbf{\Omega}'}{\partial\xi_i}\right) \otimes \mathbf{R}, & i = m+1,\ldots,\frac{m(m+1)}{2}; \\ (\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}') \otimes \frac{\partial\mathbf{R}}{\partial\xi_i}, & i = \frac{m(m+1)}{2}+1,\ldots,\frac{m(m+1)}{2}+s. \end{cases}$$

When $i = 1,\ldots,m$ and $j = 1,\ldots,m$, we have that

$$\frac{1}{2}\mathrm{tr}(\mathbf{P}\frac{\partial\mathbf{\Sigma}}{\partial\xi_i}\mathbf{P}\frac{\partial\mathbf{\Sigma}}{\partial\xi_j}) = \frac{1}{2}\mathrm{tr}\left(\mathbf{\Lambda}^{-1}\frac{\partial\mathbf{\Lambda}}{\partial\xi_i}\mathbf{\Lambda}^{-1}\frac{\partial\mathbf{\Lambda}}{\partial\xi_j}\right)\mathrm{tr}\left(\mathbf{P}_u\mathbf{R}\mathbf{P}_u\mathbf{R}\right) = \begin{cases} 0, & i \neq j; \\ \frac{n-p}{2\lambda_i^2}, & i = j. \end{cases}$$

When $i = 1,\ldots,m$ and $j = m+1,\ldots,\frac{m(m+1)}{2}$, we have that

$$\frac{1}{2}\mathrm{tr}(\mathbf{P}\frac{\partial\mathbf{\Sigma}}{\partial\xi_i}\mathbf{P}\frac{\partial\mathbf{\Sigma}}{\partial\xi_j}) = \frac{n-p}{2}\left[\mathrm{tr}\left(\frac{\partial\mathbf{\Lambda}}{\partial\xi_i}\mathbf{\Lambda}^{-1}\left\{\mathbf{\Omega}'\frac{\partial\mathbf{\Omega}}{\partial\xi_j} + \frac{\partial\mathbf{\Omega}'}{\partial\xi_j}\mathbf{\Omega}\right\}\right)\right] = 0$$

since $\mathbf{\Omega}'\frac{\partial\mathbf{\Omega}}{\partial\xi_j} + \frac{\partial\mathbf{\Omega}'}{\partial\xi_j}\mathbf{\Omega} = \frac{\partial(\mathbf{\Omega}'\mathbf{\Omega})}{\partial\xi_j} = \frac{\partial\mathbf{I}}{\partial\xi_j} = 0$. Similar calculations yield

$$\frac{1}{2}\mathrm{tr}(\mathbf{P}\frac{\partial\mathbf{\Sigma}}{\partial\xi_i}\mathbf{P}\frac{\partial\mathbf{\Sigma}}{\partial\xi_j}) = \begin{cases} \frac{1}{2\lambda_i}\mathrm{tr}\left(\mathbf{P}_u\frac{\partial\mathbf{R}}{\partial\xi_j}\right) & i = 1,\ldots,m;\ j = \frac{m(m+1)}{2}+1,\ldots,\frac{m(m+1)}{2}+s \\ \frac{n-p}{2}k_{ij} & i,j = m+1,\ldots,\frac{m(m+1)}{2} \\ 0 & i = m+1,\ldots,\frac{m(m+1)}{2};\ j = \frac{m(m+1)}{2}+1,\ldots,\frac{m(m+1)}{2}+s \\ \frac{m}{2}\mathrm{tr}\left(\mathbf{P}_u\frac{\partial\mathbf{R}}{\partial\xi_i}\mathbf{P}_u\frac{\partial\mathbf{R}}{\partial\xi_j}\right) & i,j = \frac{m(m+1)}{2}+1,\ldots,\frac{m(m+1)}{2}+s \end{cases}$$

where $k_{ij} = \mathrm{tr}\left[\mathbf{\Omega}\mathbf{\Lambda}^{-1}\mathbf{\Omega}'\left(\frac{\partial\mathbf{\Omega}}{\partial\xi_i}\mathbf{\Lambda}\mathbf{\Omega}' + \mathbf{\Omega}\mathbf{\Lambda}\frac{\partial\mathbf{\Omega}'}{\partial\xi_i}\right)\mathbf{\Omega}\mathbf{\Lambda}^{-1}\mathbf{\Omega}'\left(\frac{\partial\mathbf{\Omega}}{\partial\xi_j}\mathbf{\Lambda}\mathbf{\Omega}' + \mathbf{\Omega}\mathbf{\Lambda}\frac{\partial\mathbf{\Omega}'}{\partial\xi_j}\right)\right]$ and we have used the identity $\mathbf{P}_u\mathbf{R}\mathbf{P}_u = \mathbf{P}_u$.

Hence

$$\mathbf{I}(\boldsymbol{\xi}, D) = \begin{pmatrix} \mathbf{I}_{11} & \mathbf{0} & \mathbf{I}_{13} \\ \mathbf{0} & \mathbf{I}_{22} & \mathbf{0} \\ \mathbf{I}_{13}' & \mathbf{0} & \mathbf{I}_{33} \end{pmatrix}$$

where $\mathbf{I}_{11}$ is the $m \times m$ diagonal matrix whose ith diagonal element is $\frac{n-p}{2\lambda_i^2}$, $\mathbf{I}_{13}$ is the $m \times s$ matrix whose (i,j)th element is given by $\frac{1}{2\lambda_i}\mathrm{tr}\left(\mathbf{P}_u\frac{\partial\mathbf{R}}{\partial\xi_{\frac{m(m+1)}{2}+j}}\right)$, $\mathbf{I}_{22}$ is the $\frac{m(m-1)}{2} \times \frac{m(m-1)}{2}$ matrix whose (i,j)th element is $\left(\frac{n-p}{2}\right)k_{ij}$, and $\mathbf{I}_{33}$ is the $s \times s$ matrix whose (i,j)th element is given by $\frac{m}{2}\mathrm{tr}\left(\mathbf{P}_u\frac{\partial\mathbf{R}}{\partial\xi_{\frac{m(m+1)}{2}+i}}\mathbf{P}_u\frac{\partial\mathbf{R}}{\partial\xi_{\frac{m(m+1)}{2}+j}}\right)$. Observe that $\mathbf{I}_{11}$ and $\mathbf{I}_{22}$ do not depend on the design D. Thus we may write, using well-known results on determinants of partitioned matrices,

$$|\mathbf{I}(\boldsymbol{\xi}, D)| = |\mathbf{I}_{11}||\mathbf{I}_{22}||\mathbf{I}_{33} - \mathbf{I}_{13}'\mathbf{I}_{11}^{-1}\mathbf{I}_{13}|.$$

It is easily verified that $\mathbf{I}_{13}'\mathbf{I}_{11}^{-1}\mathbf{I}_{13}$ is an $s \times s$ matrix whose (i, j)th element is given by

$$\frac{m}{2(n-p)}\mathrm{tr}\left(\mathbf{P}_u \frac{\partial \mathbf{R}}{\partial \xi_{\frac{m(m+1)}{2}+i}}\right)\mathrm{tr}\left(\mathbf{P}_u \frac{\partial \mathbf{R}}{\partial \xi_{\frac{m(m+1)}{2}+j}}\right).$$

Hence,

$$|\mathbf{I}(\boldsymbol{\xi}, D)| = \left(\frac{2m\sigma_1^4}{n-p}\right)|\mathbf{I}_{11}||\mathbf{I}_{22}||\mathbf{I}_u(\sigma_1^2, \boldsymbol{\rho}, D)|,$$

where $\mathbf{I}_u(\sigma_1^2, \boldsymbol{\rho}, D)$ is the REML information matrix for the corresponding univariate design problem. It follows that minimizing $1/|\mathbf{I}(\boldsymbol{\xi}, D)|$ with respect to D is equivalent to minimizing its univariate counterpart $1/|\mathbf{I}_u(\sigma_1^2, \boldsymbol{\rho}, D)|$. This completes the proof of part (b).