VERSITA

TATRA MOUNTAINS
Mathematical Publications
DOI: 10.2478/v10127-012-0016-x
Tatra Mt. Math. Publ. **51** (2012), 161–171

NETWORK-RELATED PROBLEMS IN OPTIMAL EXPERIMENTAL DESIGN AND SECOND ORDER CONE PROGRAMMING

Guillaume Sagnol

ABSTRACT. In the past few years several applications of optimal experimental designs have emerged to optimize the measurements in communication networks. The optimal design problems arising from this kind of applications share three interesting properties: (i) measurements are only available at a small number of locations of the network; (ii) each monitor can simultaneously measure several quantities, which can be modeled by "multiresponse experiments"; (iii) the observation matrices depend on the topology of the network. We give an overview of these experimental design problems and recall recent results for the computation of optimal designs by Second Order Cone Programming (SOCP). New results for the network-monitoring of a discrete time process are presented. In particular, we show that the optimal design problem for the monitoring of an AR1 process can be reduced to the standard form and we give experimental results.

1. Network monitoring and optimal design

The approximate theory for the optimal design of experiments is an important branch of statistics, and we refer the reader to the monograph of Pukelsheim [5] for a comprehensive review on the subject. In the classical version of the problem, an experimenter wants to estimate a vector of unknown parameters $\boldsymbol{\theta} \in \mathbb{R}^m$. To this end, he has s experiments available. The ith experiment yields a measurement $\boldsymbol{y_i} \in \mathbb{R}^{l_i}$ such that

$$\mathbb{E}[\boldsymbol{y_i}] = X_i\boldsymbol{\theta} \quad \text{and} \quad \mathrm{Var}[\boldsymbol{y_i}] = \Sigma_i.$$

Experiments are uncorrelated, i.e.,

$$\mathbb{E}[\boldsymbol{y_i}\boldsymbol{y_j}^T] = 0 \qquad \text{for } i \neq j.$$

2010 Mathematics Subject Classification: 62K05.
Keywords: optimal design, network monitoring, AR1 process, Second Order Cone Programming.

The matrix $X_i \in \mathbb{R}^{l_i \times m}$ is called the observation matrix of the ith experiment. Given a matrix of coefficients $K \in \mathbb{R}^{m\times k}$, the goal is to decide the fraction w_i of the optimal effort to allocate to the ith experiment (for all $i \in [s] := \{1,\ldots,s\}$), so as to estimate the vector $K^T\boldsymbol{\theta} \in \mathbb{R}^k$ with the best possible accuracy. The vector $\boldsymbol{w} \in \mathbb{R}^s_+$ sums to 1 and is called a design.

We define the information matrix of the design $\boldsymbol{w}$ by

$$M(\boldsymbol{w}) := \sum_{i=1}^{s} w_i X_i^T \Sigma_i^{-1} X_i.$$

The standard approach is to minimize, with respect to the design variable $\boldsymbol{w}$, a spectral information function Φ which is applied to the information matrix:

$$\min_{\boldsymbol{w}} \left\{ \Phi\big(M(\boldsymbol{w})\big) : \boldsymbol{w} > \mathbf{0}, \sum_{i=1}^{s} w_i = 1 \right\}.$$

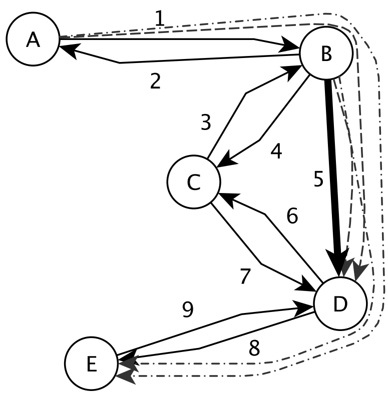
Common choices for the function Φ are the criterions of $E-$optimality (E stands for Eigenvalue), $A-$optimality (for Average), and $D-$optimality (for Determinant), defined respectively on the set of positive semidefinite matrices Z whose range includes the columns of K by (see [5, Chapter 6]):

$$\begin{aligned}
\Phi_K^{(E)}(Z) &= \lambda_{\max}(K^T Z^\dagger K),\\
\Phi_K^{(A)}(Z) &= \operatorname{trace} K^T Z^\dagger K,\\
\Phi_K^{(D)}(Z) &= \det K^T Z^\dagger K,
\end{aligned}$$

where $M^\dagger$ is the Moore-Penrose pseudo-inverse of M and $\lambda_{\max}$ denotes the largest eigenvalue.

In most common situations, the number m of unknown parameters is rather small ($m \leq 10$), while the number s of possible experiments is very large, typically coming from the discretization of a multidimensional, compact set $\mathcal{X}$. Recently however, a new class of instances has arisen from network monitoring problems. Here, the number s of experiments remains reasonably small (in the order of the number of nodes in the graph), while the number m of unknown parameters is very large ($m = O(s^2)$). The problem is not ill-posed despite the relation $s \leq m$ (less experiments than unknowns), because the experiments are *multiresponse*, meaning that the ith experiment yields $l_i > 1$ simultaneous observations.

These special instances of optimal experimental design appear when the experimenter wants to set monitors over a network, in order to estimate certain quantities such as a performance indicator of an Internet network [10], or the volume of each origin-destination flows [7], [8]. A common point to these network--related problems is that the observation matrices only depend on the topology and the routing of the flows in the graph, and the variance matrices depend on a prior estimate of the flow volumes. We give below a toy-example showing how to construct a network-monitoring optimal design problem. Then, we will review in Section 2 recent results based on Second Order Cone Programming (SOCP) for the computation of these optimal designs. Finally, we will show in Section 3 that the optimal monitoring of a discrete time process over a network can also be formulated under the standard form, and we show some experimental results.



Optimal designs on the edges for $K = I$.

Edge	1	2	3	4	5	6	7	8	9
E-optimal	11.1	11.1	16.6	5.5	11.1	16.6	5.5	11.1	11.1
A-optimal	11.2	11.0	14.5	7.3	11.5	14.5	7.3	11.2	11.0
D-optimal	11.0	10.8	14.1	7.9	11.9	14.1	7.9	11.0	10.8

FIGURE 1. *Left*: Graph of a toy example with 5 nodes and 9 edges. The dashed lines represent the 4 flows traversing edge 5 ($A \to D$, $B \to D$, $A \to E$, $B \to E$). *Right*: weights of the $E-$, $A-$, and $D-$optimal designs for this network (in %), for the observation matrices X_i of the subvector observation model and the variance matrices Σ_i of the toll evasion model described in the text, when K is the identity matrix.

We consider the simple network represented in Figure 1. The unknown parameter $\boldsymbol{\theta}$ can be any measurable feature of the origin-destination (OD) flows, for example the volume of each flow during a given period of time, the average length of the individuals in each flow (e.g., length of Internet packets or vehicles), or the proportion of each flow belonging to a special category (e.g., fraction of FTP traffic). The experimental effort should be distributed over monitors located on the edges of the network (in a computer network, the monitors correspond to a measurement software, while in a road network they can be human pollers who stop the vehicles).

Now we shall distinguish two cases. The first case appears when the monitors are able to infer the source and the destination of the sampled individuals. For all edge e in the network, this results in an observation of every OD pair whose path includes e (for simplicity, we assume here that every individual chooses the shortest path from his source to his destination). For example, a monitor on edge 5 observes the four flows which traverse it,

$$\theta_{A\to D},\ \theta_{B\to D},\ \theta_{A\to E}, \quad \text{and} \quad \theta_{B\to E}.$$

Hence the observation matrix X_5 has 4 rows (one per observed flow) and $5\times 4 = 20$ columns (one per OD pair)

$$X_5 = \begin{array}{c} \begin{array}{cccccccc} \rotatebox{90}{A$\to$B} & \rotatebox{90}{A$\to$C} & \rotatebox{90}{A$\to$D} & \rotatebox{90}{A$\to$E} & \cdots & \rotatebox{90}{B$\to$D} & \rotatebox{90}{B$\to$E} & \cdots \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots \end{pmatrix} \end{array}.$$

We obtain in a similar manner an observation matrix for each edge if the flows traversing edge e are $j_1, \ldots, j_{l_e}$, the entry (k, j_k) of X_e is a 1 for all $k \in [l_e]$, the other entries are 0. We call this case the *subvector observation* model, because each experiment directly gives an estimate of a subvector of $\boldsymbol{\theta}$.

In the second case, which we shall call the *destination only* model, the monitors are only able to infer the destination of the sampled individuals (and not their origin—which makes sense on an computer network, because the packets do not store the history of the nodes they have visited [7]). In this situation, we obtain, for all edge e, an observation for every destination D that is reachable from e, by summing the desired feature for the OD pairs that traverse e and have the destination D. Practically, the new observation matrix X'_e is obtained by summing the rows of X_e that correspond to flows having a common destination. In the example above, since the flows traversing edge 5 go to either D or E, we can measure two quantities on this edge, namely

$$(\theta_{A\to D} + \theta_{B\to D}) \quad \text{and} \quad (\theta_{A\to E} + \theta_{B\to E}),$$

$$X'_5 = \begin{array}{c} \begin{array}{cccccccc} \rotatebox{90}{A$\to$B} & \rotatebox{90}{A$\to$C} & \rotatebox{90}{A$\to$D} & \rotatebox{90}{A$\to$E} & \cdots & \rotatebox{90}{B$\to$D} & \rotatebox{90}{B$\to$E} & \cdots \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 & \cdots & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots & 0 & 1 & \cdots \end{pmatrix} \end{array}.$$

The variance-covariance matrix Σ_e depends on the sampling scheme and the feature of interest which the experimenter wants to estimate. An example with a randomized sampling to measure Internet flow volumes can be found in [7].

Now we give another example which is motivated by an application to toll enforcement on German motorways [1]. Here, the control tours of inspectors should be optimized, in order to make the the toll enforcement more efficient. We propose to compute a design (i.e., an allocation of the control density to the road segments of the network) which leads to the most accurate estimation of the number of toll evaders on each OD pair. In a follow-up work, we want to use this design as a target for the integer program described in [1].

We consider an edge e traversed by l OD pairs, which we denote by $1, \ldots, l$ for simplicity. The volume of traffic x_j on each OD pair is known, as well as the total volume of traffic y_e traversing edge e (during a given period of time). Let κ denote the number of trucks that an inspector can control. We denote by $\boldsymbol{N}^+ \in \mathbb{R}^l$ (resp. $\boldsymbol{N}^- \in \mathbb{R}^l$) the vector of counts of toll evaders (resp. payers) from the OD pairs $1, \ldots, l$ checked by the inspector (this is the subvector observation model described above, because the inspector knows the origin and the destination of the trucks he checks). The joint distribution of $(\boldsymbol{N}^+, \boldsymbol{N}^-)$ is multinomial (modulo the standard approximation of the hypergeometric law by a multinomial), with κ trials and proportions

$$\left[\left(\frac{x_j p_j}{y_e}\right)_{j=1,\ldots,l}, \left(\frac{x_j(1-p_j)}{y_e}\right)_{j=1,\ldots,l}\right],$$

where p_j is the proportion of evaders on the jth OD pair. The Cramer Rao bound shows that any unbiased estimator for the number of evaders

$$\theta_j = x_j p_j \text{ (for } j = 1, \ldots, l), \text{ satisfies } \operatorname{Var}[\hat{\boldsymbol{\theta}}] \succeq V_e$$

(in the Löwner ordering sense), where V_e is the $l \times l-$diagonal matrix with elements $\kappa^{-1} x_j y_e p_j (1-p_j)$. Moreover, the unbiased estimators for θ_j are of the form

$$\hat{\theta}_j = \frac{y_e}{\kappa}\left((1-\alpha_j) N_j^+ + \alpha_j \left(\kappa \frac{x_j}{y_e} - N_j^-\right)\right),$$

and the lower bound for $\operatorname{Var}[\hat{\boldsymbol{\theta}}]$ is attained for $\alpha_j = p_j$. In practice, p_j is not known and we replace it by a prior estimate $\hat{p}_j$, i.e., we set Σ_e to the diagonal matrix with the elements

$$\kappa^{-1} x_j y_e \hat{p}_j (1-\hat{p}_j) \quad (\text{for } \ j = 1, \ldots, l).$$

In the toy example of Figure 1, the traffic on edge 5 satisfies the relation $y_5 = x_{A\to D} + x_{B\to D} + x_{A\to E} + x_{B\to E}$, and if we take the same prior $\hat{p}_j$ for the evasion rate on every OD pair, the matrix Σ_5 will be proportional to

$$\operatorname{Diag}(y_5 x_{A\to D}, y_5 x_{A\to E}, y_5 x_{B\to D}, y_5 x_{B\to E}).$$

The optimal designs indicated on Figure 1 were computed for a traffic set to $x_j = 1$ on every OD pair.

2. The Second-Order Cone Programming approach

We have shown in [6] that an $A-$optimal design for $K^T\boldsymbol{\theta}$, i.e., a design which solves

$$\min_{\boldsymbol{w}} \left\{ \phi_K(\boldsymbol{w}) := \operatorname{trace} K^T \left(\sum_{i\in[s]} w_i X_i^T \Sigma_i^{-1} X_i \right)^{\dagger} K : \boldsymbol{w} \geq \mathbf{0},\ \sum_{i=1}^{s} w_i = 1 \right\}, \quad (1)$$

with the additional implicit constraint $\operatorname{Range}\big(M(\boldsymbol{w})\big) \supseteq \operatorname{Range}(K)$, is obtained by normalizing any vector $\boldsymbol{\mu^*}$ which solves the following Second Order Cone Program (SOCP)

$$\min_{\substack{\boldsymbol{\mu}\in\mathbb{R}^s\\ H_i\in\mathbb{R}^{l_i\times k}}} \sum_i \mu_i \quad (2)$$

such that

$$K = \sum_i X_i^T \Sigma_i^{-\frac{1}{2}} H_i, \quad \|H_i\|_F \leq \mu_i \quad (\text{for all } i \in [s]).$$

In this optimization problem, $\|M\|_F := \left(\sum_{i,j} M_{i,j}^2\right)^{1/2}$ denotes the Frobenius norm of M. SOCP is a general class of optimization problems [4] which can be solved efficiently by interior point codes such as SeDuMi [11]. When the dimension of the problem becomes large (typically $m \geq 10^4$, which occurs in network monitoring problems), classical algorithms fail to be efficient: Wynn–Fedorov exchange algorithm [13] has a slow convergence, and multiplicative update algorithms [12] require a prohibitive full-rank update of the Cholesky factorization of an $m\times m-$matrix at each iteration. In contrast, computational results of [6] show that SOCP solvers perform well as long as the observation matrices are sparse (this is the case indeed for network problems) and the number of columns k of K is small (in particular, for $\boldsymbol{c}-$optimality, where $k = 1$ and $K = \boldsymbol{c}$ is a column vector).

If both k and m are large however (in particular when the full parameter $\boldsymbol{\theta}$ is of interest, i.e., $K = I$), no tractable algorithm is known to compute an $A-$optimal design. (Except in the easy case where the information matrix $M(\boldsymbol{w})$ is diagonal, which happens in the subvector observation model.) Based on the observation that $\mathbb{E}\big[\phi_{\boldsymbol{c}}(\boldsymbol{w})\big] = \phi_K(\boldsymbol{w})$, where the expectation is taken with respect to $\boldsymbol{c} \sim \mathcal{N}(\mathbf{0}, KK^T)$, a heuristic approach based on $\boldsymbol{c}-$optimality was proposed in [7] to find a design $\boldsymbol{w}$ such that $\phi(\boldsymbol{w})$ approximates the optimal value of Problem (1). Some vectors $\boldsymbol{c_1}, \ldots, \boldsymbol{c_N}$ are generated from the distribution $\mathcal{N}(\mathbf{0}, KK^T)$, and we take the mean of the corresponding $\boldsymbol{c_i}-$optimal designs. In the latter article, the design found with this procedure is essentially tested for its performance with respect to the applied networking problem. To go further, we have tested this heuristic design on several networks from topology-zoo.org [2]. The first column of Table 1 describes the instance: name of the network, type of information that the monitor can read from the flows ('OD' for the subvector observation model, 'D' if only the destination of the flows can be inferred), and location of the measurements ('links' or 'nodes'). The four next columns describe the size of the instance. The last two columns give the L_1-relative error $\frac{\|\boldsymbol{w}-\boldsymbol{w_A}\|_1}{\|\boldsymbol{w_A}\|_1}$ and the $A-$efficiency $\frac{\phi_I(\boldsymbol{w_A})}{\phi_I(\boldsymbol{w})}$ computed by comparing the average $\boldsymbol{w}$ of $N = 100$ $\boldsymbol{c}-$optimal designs and the true $A-$optimal design $\boldsymbol{w_A}$. Remarkably, although the design found by this technique is not always very close to the $A-$optimal design, the A-efficiency is excellent for every instance. The reasons of the good behaviour of this heuristic for network-monitoring optimal design problems are still unknown. But we think that the present analysis justifies the use of this technique when no tractable algorithm is available.

An independent SOCP formulation was discovered by Singhal and Michailidis for a particular optimal design problem in a filtering context [9]. Here, the unknown parameter is observed over time, and the process $\boldsymbol{\theta_t} (t \in \mathbb{N})$ is assumed to be a random walk $\boldsymbol{\theta_{t+1}} = \boldsymbol{\theta_t} + \boldsymbol{\epsilon_t}$, where the noise vectors $\boldsymbol{\epsilon_1}, \boldsymbol{\epsilon_2}, \ldots$ are i.i.d., centered, and have a known diagonal covariance matrix Q. The authors further assume that each X_i has only one nonzero per row (subvector observation model) and each Σ_i is diagonal. They give an SOCP to compute a steady-state $E-$optimal design, i.e., a design which maximizes the smallest eigenvalue of the asymptotic information matrix in the steady state of the Kalman filter. In the next section, we also study a network monitoring problem over time, but we want to allocate in advance the monitoring resource for one day of measurements, when the experimental effort is not required to be spread uniformly during the day.

3. Optimal monitoring of a discrete time process

Most network flows exhibit strong diurnal patterns [3]. Therefore, it makes sense to distribute the experimental effort not only over different locations of the networks, but also over time. Typically, the goal is to save monitoring resources at night in order to use them at day, when the traffic is more important. If we divide the day in T periods $t = 1, \ldots, T$, then a straightforward approach is to consider the augmented vector of unknown parameters $\boldsymbol{\theta} = [\boldsymbol{\theta_1}^T, \ldots, \boldsymbol{\theta_T}^T]^T$, so that the observation equations

$$\boldsymbol{y_{i,t}} = X_i \boldsymbol{\theta_t} + \boldsymbol{\epsilon_{i,t}}, \quad \mathrm{Var}[\boldsymbol{\epsilon_{i,t}}] = \Sigma_{i,t} \quad \text{(for all } i, t \in [s] \times [T])$$

may be rewritten in the standard form

$$\boldsymbol{y_{i,t}} = [\ \overset{(\boldsymbol{\theta_1})}{0}\ , \ldots, \overset{(\boldsymbol{\theta_t})}{X_i}, \ldots, \overset{(\boldsymbol{\theta_T})}{0}\ ]\, \boldsymbol{\theta} + \boldsymbol{\varepsilon_{i,t}}, \quad \mathrm{Var}[\boldsymbol{\varepsilon_{i,t}}] = \Sigma_{i,t}. \tag{3}$$

Table 1. $A-$efficiency and L_1-relative error for the design found by averaging $N = 100$ $\boldsymbol{c}-$optimal designs ; exp, obs, par and nnz indicate respectively the number s of experiments, the total number $\sum_{i=1}^{s} l_i$ of observations, the number m of parameters, and the total number of nonzeroes in $X_1, \ldots, X_s$.

Instance	Exp.	Obs.	Par.	NNZ	L1 error (%)	A-efficiency (%)
Abilene_OD_links	51	743	121	1386	12.1	99.212
Abilene_D_links	50	285	121	693	6.0	99.242
Abilene_OD_nodes	12	501	121	1144	6.3	99.812
Abilene_D_nodes	11	274	121	572	2.5	99.928
Bellcanada_OD_links	225	29987	2304	59526	10.3	99.664
Bellcanada_D_links	224	5562	2304	29763	4.1	99.911
Bellcanada_OD_nodes	49	25379	2304	54918	5.6	99.827
Bellcanada_D_nodes	48	5514	2304	27459	3.1	99.797
Dfn_OD_links	291	27287	3364	53994	9.9	99.890
Dfn_D_links	290	8386	3364	26997	4.5	99.953
Dfn_OD_nodes	59	20559	3364	47266	3.9	99.846
Dfn_D_nodes	58	8328	3364	23633	4.8	99.655
Garr201012_OD_links	253	26730	3136	52956	9.4	99.975
Garr201012_D_links	252	7159	3136	26478	3.7	99.947
Garr201012_OD_nodes	57	20458	3136	46684	5.0	99.741
Garr201012_D_nodes	56	7103	3136	23342	6.6	99.351
Geant2010_OD_links	187	13385	1369	26398	8.8	99.734
Geant2010_D_links	186	3403	1369	13199	3.8	99.864
Geant2010_OD_nodes	38	10647	1369	23660	5.2	99.753
Geant2010_D_nodes	37	3366	1369	11830	4.7	99.673
Gtspoland_OD_links	141	9904	1089	19528	13.2	99.494
Gtspoland_D_links	140	2425	1089	9764	4.8	99.847
Gtspoland_OD_nodes	34	7726	1089	17350	5.5	99.840
Gtspoland_D_nodes	33	2392	1089	8675	3.8	99.650
Renater2010_OD_links	171	13384	1369	26428	12.2	99.611
Renater2010_D_links	170	3267	1369	13214	4.9	99.875
Renater2010_OD_nodes	38	10646	1369	23690	6.0	99.748
Renater2010_D_nodes	37	3230	1369	11845	2.8	99.891

Note that in this model, the dependence on time of the observation equations only appears through the variance-covariance matrices $\Sigma_{i,t}$ of the observations which usually depend on the flow volumes at location i and time t. Also note that we have implicitly assumed that the measurements are mutually independent, i.e., $\mathbb{E}[\boldsymbol{\varepsilon_{i,t}}\boldsymbol{\varepsilon_{j,\tau}}^T] = 0$ whenever $(i,t) \neq (j,\tau)$, because the noise is only due to sampling effects.

In the model relying on equation (3), we ignore the fact that $\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \ldots$ is a structured time process, and that the measurement of $\boldsymbol{\theta_t}$ potentially carries some information about the parameter at other time periods. To tackle this issue, we propose to study an optimal design problem with a simple time model of the flows, based on an autoregressive process of the first order: for all $t \in [T]$,

$$\boldsymbol{\theta_t} = \bar{\boldsymbol{\theta}} + \boldsymbol{d_t}, \tag{4}$$

$$\boldsymbol{d_{t+1}} = P\boldsymbol{d_t} + \boldsymbol{\eta_t}, \tag{5}$$

where $\bar{\boldsymbol{\theta}}$ is the mean of the process, P is a coefficient matrix with a spectral norm smaller than 1, and the process $\boldsymbol{\eta_1}, \boldsymbol{\eta_2}, \ldots$ is i.i.d. with

$$\mathbb{E}[\boldsymbol{\eta_t}] = \mathbf{0} \quad \text{and} \quad \operatorname{Var}[\boldsymbol{\eta_t}] = Q.$$

If we define the augmented parameter

$$\tilde{\boldsymbol{\theta}} = \left[\bar{\boldsymbol{\theta}}^T, \boldsymbol{d_1}^T, \ldots, \boldsymbol{d_T}^T\right]^T,$$

the observation equations can be written

$$\boldsymbol{y_{i,t}} = [\overset{(\bar{\boldsymbol{\theta}})}{X_i},\ \overset{(\boldsymbol{d_1})}{0}, \ldots, \overset{(\boldsymbol{d_t})}{X_i}, \ldots, \overset{(\boldsymbol{d_T})}{0}\ ]\, \tilde{\boldsymbol{\theta}} + \boldsymbol{\varepsilon_{i,t}}, \quad \operatorname{Var}[\boldsymbol{\varepsilon_{i,t}}] = \Sigma_{i,t} \tag{6}$$
$$\text{(for all } (i,t) \in [s] \times [T]).$$

We also note that the autoregression equation (5) are equivalent to a collection of virtual prior observations

$$\boldsymbol{z_t}(= \mathbf{0}) = [\ \overset{(\bar{\boldsymbol{\theta}})}{0},\ \overset{(\boldsymbol{d_1})}{0}, \ldots, \overset{(\boldsymbol{d_t})}{P},\ \overset{(\boldsymbol{d_{t+1}})}{-I}, \ldots, \overset{(\boldsymbol{d_T})}{0}\ ]\, \tilde{\boldsymbol{\theta}} + \boldsymbol{\eta_t}, \quad \operatorname{Var}[\boldsymbol{\eta_t}] = Q. \tag{7}$$
$$\text{(for all } t \in [T-1]).$$

Finally, the full parameter is $\boldsymbol{\theta} = K^T \tilde{\boldsymbol{\theta}}$, where

$$K^T = \begin{pmatrix} I & I & 0 & \cdots & 0 \\ I & 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ I & 0 & 0 & \cdots & I \end{pmatrix},$$

and an $A-$optimal design for $K^T\boldsymbol{\theta}$ can be found by SOCP or multiplicative algorithms (note that we must use an SOCP whose form differs from (2), because of the prior observations in equation (7), see [6]).

We have computed the $A-$optimal designs associated to the observation models (3) and (6)–(7), for the application to toll enforcement described in Section 1 on the motorways of the region of Berlin-Brandenburg in Germany ($s = 90$ edges, $m = 494$ pairs). In the latter model we have set

$$P = 0.5I \quad \text{and} \quad Q^{1/2} = 0.025\,\mathrm{Diag}(\bar{\boldsymbol{x}}),$$

where $\bar{x}_j$ is the average flow volume on the jth OD.

For this experiment the day was divided in 8 time slots. The temporal components of the optimal designs (i.e., $w_t = \sum_{i=1}^{s} w_{i,t}$) are plotted in Figure 2, together with the diurnal evolution of the traffic volume. While the design for Model (3) follows the daily trend of the traffic, we see that the design based on an AR1 process is mainly concentrated during the night, where flows are less important and measurements are more accurate. Nevertheless, we point out that assuming a time structure might not be adapted for this toll enforcement problem, because the estimation of the number of toll evaders should rely on actual controls rather than sophisticated time models.

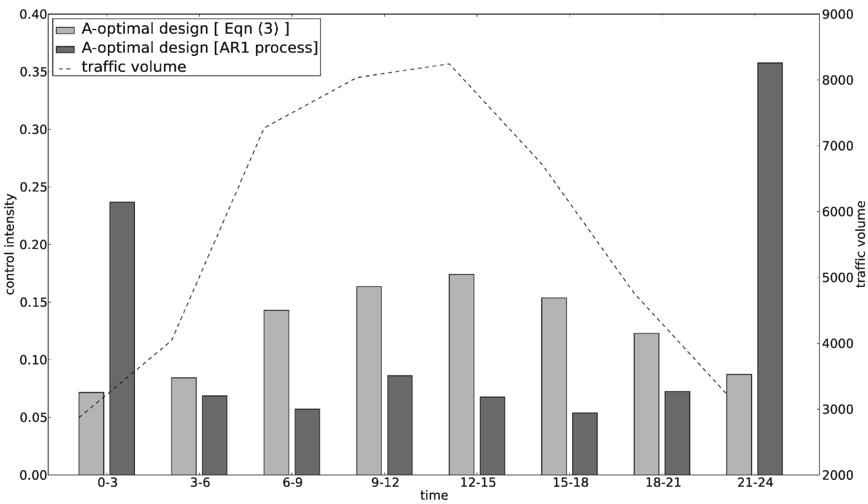


Figure 2. Optimal designs and total traffic volume vs. time in the region of Berlin Brandenburg, for the models based on independent measurements (3) and an AR1 process (6)–(7).

REFERENCES

[1] BORNDÖRFER, R.—SAGNOL, G.—SWARAT, E.: *An IP approach to toll enforcement optimization on german motorways*, Tech. Rep. ZIB, Report 11–42, Zuse Institut Berlin, 2011.

[2] KNIGHT, S.—NGUYEN, H. X.—FALKNER, N.—BOWDEN, R.—ROUGHAN, M.: *The internet topology zoo,* IEEE J. Select. Areas Commun. **29** (2011), 1765–1775.

[3] LAKHINA, A.—PAPAGIANNAKI, K.—CROVELLA, M.—DIOT, C.—KOLACZYK, E.–TAFT, N.: *Structural analysis of network traffic flows,* in: Proc. ACM SIGMETRICS '04 (E. G. Coffman Jr., Z. Liu and A. Merchant, eds.), New York, NY, SIGMETRICS Perform. Eval. Rev., Vol. 32, ACM New York, NY, USA, 2004, pp. 61–72.

[4] LOBO, M. S.—VANDENBERGHE, L.—BOYD, S.—LEBRET, H.: *Applications of second-order cone programming,* Linear Algebra Appl. **284** (1998), 193–228.

[5] PUKELSHEIM, F.: *Optimal Design of Experiments*. Wiley, New York, 1993.

[6] SAGNOL, G.: *Computing optimal designs of multiresponse experiments reduces to second-order cone programming,* J. Statist. Plann. Inference **141** (2011), 1684–1708.

[7] SAGNOL, G.—GAUBERT, S.—BOUHTOU, M.: *Optimal monitoring on large networks by successive c-optimal designs,* in: Proc. of the 22nd Internat. Teletraffic Congr.—ITC 22, Amsterdam, The Netherlands, IEEE, Piscataway, N. J., 2010, pp. 1–8.

[8] SINGHAL, H.—MICHAILIDIS, G.: *Optimal sampling in state space models with applications to network monitoring,* in: Proc. ACM SIGMETRICS '08, Annapolis, Maryland, USA, SIGMETRICS Perform. Eval. Rev., Vol. 36, ACM New York, NY, USA, 2008, pp. 145–156.

[9] SINGHAL, H.—MICHAILIDIS, G.: *Optimal experiment design in a filtering context with application to sampled network data,* Ann. Appl. Stat. **4** (2010), 78–93.

[10] SONG, H. H.—QIU, L.—ZHANG, Y.: *Netquest: A flexible framework for largescale network measurement,* in: ACM SIGMETRICS '06 (R. A. Marie, P. B. Key and E. Smirni, eds.), St. Malo, France, 2006, SIGMETRICS Perform. Eval. Rev., Vol. 34, ACM New York, NY, USA, 2006, pp. 121–132.

[11] STURM, J. F.: *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones,* Optim. Methods Softw. **11–12** (1999), 625–653.

[12] TITTERINGTON, D. M.: *Algorithms for computing D-optimal design on finite design spaces,* in: Proc. of the 1976 Conf. on Information Science and Systems, Baltimore, USA, 1976, Dept. of Electronic Engineering, John Hopkins University, Baltimore, USA, 1976, pp. 213–216.

[13] WYNN, H. P.: *The sequential generation of D-optimum experimental designs,* Ann. Math. Stat. **41** (1970), 1655–1664.

Received March 22, 2012

Zuse Institut Berlin
Takustr. 7
D–14195 Berlin
GERMANY
E-mail: sagnol@zib.de