



**HAL**  
open science

# Covariate-adjusted response-adaptive compromise optimal designs

Asya Metelkina, Luc Pronzato

► **To cite this version:**

Asya Metelkina, Luc Pronzato. Covariate-adjusted response-adaptive compromise optimal designs. 2015. hal-01246252v1

**HAL Id: hal-01246252**

**<https://hal.science/hal-01246252v1>**

Preprint submitted on 18 Dec 2015 (v1), last revised 2 Nov 2017 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Covariate-adjusted response-adaptive compromise optimal designs\*

Asya Metelkina and Luc Pronzato

*Laboratoire I3S, UNS-CNRS  
Bât. Euclide, Les Algorithmes, BP 121  
2000 route des Lucioles  
06903 Sophia Antipolis cedex, France  
e-mail: metelkin@unice.fr, e-mail: pronzato@cnrs.fr*

**Abstract:** Covariate-adaptive treatment allocation is considered in the situation when a compromise must be made between information (about the dependency of the probability of success of each treatment upon influential covariates) and cost (in terms of number of subjects receiving the poorest treatment). Information is measured through a design criterion for parameter estimation, the cost is additive and is related to the success probabilities. Within the framework of approximate design theory, the determination of optimal allocations forms a compound design problem. We show that when the covariates are i.i.d. with a probability measure  $\mu$ , its solution possesses strong similarities with the construction of optimal design measures bounded by  $\mu$ . We characterize optimal designs through an Equivalence Theorem and construct a covariate-adaptive sequential allocation strategy that converges to the optimum. Our new optimal designs can be used as benchmarks for other, more usual, allocation methods. A response-adaptive implementation is possible for practical applications with unknown model parameters. Several illustrative examples are provided.

**MSC 2010 subject classifications:** Primary 62K05; secondary 62P10.

**Keywords and phrases:** optimal design, treatment allocation, bounded design measure, equivalence theorem, multi-armed bandit problem.

## 1. Introduction and motivation

We consider a treatment allocation problem with  $K$  treatments for which the probabilities of success depend on side information given by covariates. The response  $Y_t = Y_t(X)$  of a subject with covariates  $X$  to treatment  $t$  satisfies

$$\mathbb{E}\{Y_t|X = x, t = k\} = \eta_k(x, \theta_k), \quad k \in \{1, \dots, K\}, \quad (1.1)$$

where  $\theta_k$  denotes the (unknown) value of the model parameters for treatment  $k$  and the functions  $\eta_k$  are known. In particular, this covers the case of binary responses  $Y_t \in \{0, 1\}$ , with  $\text{Prob}[Y_t = 1|X = x, t = k] = \eta_k(x, \theta_k)$ , with logistic regression as a typical example. Throughout the paper we consider scalar

---

\*This work was initiated by discussions with several participants at the DAE (Design and Analysis of Experiments) Conference at University of Missouri, Columbia, in Oct. 2009; the support and hospitality of the organizers is gratefully acknowledged. Parts of it have been presented at the 6th and 8th International Conferences of the ERCIM WG on Computational and Methodological Statistics (ERCIM), University of London, December 2013 and 2015.

responses  $\eta_k$ , but the multivariate situation case may be considered as well, see, e.g., Dragalin and Fedorov (2006); Dragalin et al. (2008a,b); Rabie and Flournoy (2013) for bivariate binary responses corresponding to efficacy and toxicity. Note that the different models may have some parameters in common; i.e., the vectors  $\theta_k$  may share some components. We suppose that the covariates are i.i.d. among subjects, with some probability measure  $\mu$ . The responses are independent too; that is, the random vectors  $(X_i, Y_1(X), \dots, Y_k(X_i))$  are i.i.d.

In a bandit framework, this allocation problem corresponds to a multi-armed bandit problem which involves sampling from non-homogeneous populations, the response in each population being determined by a random vector of covariates and a vector of unknown parameters. The problem can be described as that of choosing between arms of a slot machine, where a random arm-dependent reward is realized each time an arm is pulled, see, e.g., Goldenshluger and Zeevi (2013).

We consider trials that aim at explicitly taking two conflicting objectives into account. The first one concerns statistical inference about the response models  $\eta_k$  and consists in estimating the  $\theta_k$  with good precision. For that reason, the allocation rule that we propose will rely on a classical criterion for experimental design, related to the precision of the estimation of the  $\theta_k$ . The second objective concerns individual ethics: one aims at minimizing a cumulative regret relative to allocation of each subject to the best treatment. The overall strategy may correspond to maximizing information under a constraint on regret, or minimize the regret with a constraint on information. From Lagrangian duality, when the information and regret functionals are respectively concave and convex (we shall use a linear regret in what follows), such strategies amount at maximizing a linear combination of information and regret. By tuning the (scalar) Lagrange coefficient, we can then set a compromise between the information gained from the trial and the efficient treatment of individuals enrolled in the trial, see Dragalin and Fedorov (2006); Dragalin et al. (2008a); Pronzato (2010).

Covariate-adaptive allocation rules are considered, for which the decision of which treatment to assign to the  $n$ -th subject is taken after observation of its covariates  $X_n$ . The main objective of the paper is to derive simple sequential allocation rules that are asymptotically optimal for the compromise criterion that we propose. The concept of design measures is natural for investigating asymptotic properties, and we shall decompose  $\mu$  into positive measures  $\xi_1, \dots, \xi_K$  on the set of covariates, with  $\xi_k$  the fraction of  $\mu$  corresponding to subjects allocated to treatment  $k$ . Any such decomposition defines a design  $\underline{\xi}$ . The presence of the constraint  $\sum_{k=1}^K \xi_k = \mu$  on  $\underline{\xi}$  introduces similarities with the construction of optimal bounded design measures, and an Equivalence Theorem is derived in Section 2, in a form similar to that in (Wynn, 1982; Fedorov, 1989; Sahn and Schwabe, 2001), see also Pronzato (2004). Optimal designs  $\underline{\xi}^*[\mu]$  for an information/regret compromise are then characterized. The introduction of randomization is considered in Section 3, with a characterization of optimal designs having a given randomization factor. Two covariate-adaptive allocation

rules are presented in Section 4. The first one relies on the prior construction of an (oracle) optimal design  $\underline{\xi}^*[\mu]$ , the second is doubly adaptive and does not assume knowledge of  $\mu$ . The paper focusses on locally optimum design, where the model parameters  $\theta_k$  are fixed to some prior nominal values. Covariate-Adjusted Response-Adaptive (CARA) rules, see Hu and Rosenberger (2006); Hu and Hu (2012); Zhang et al. (2007); Zhang and Hu (2009); Baldi Antognini and Zagoraiou (2015), where allocation of the  $n$ -th subject to one of the  $K$  treatments depends on the current  $X_n$  and estimated values  $\theta_k^{n-1}$  of the  $\theta_k$  for the  $K$  models, are briefly considered in Section 5.

## 2. Optimal covariate-adaptive design

### 2.1. Allocation criterion

Let  $\mathcal{X} \subset \mathbb{R}^d$  denote the space of covariates  $X_i$ , which are assumed to be independently identically distributed (i.i.d.) with a probability measure  $\mu$  such that  $\mu(\mathcal{X}) = 1$ . We shall consider the two following situations,

H1a:  $\mathcal{X}$  is finite,

H1b:  $\mathcal{X}$  is a compact subset of  $\mathbb{R}^d$  with non-empty interior  $\text{int}(\mathcal{X})$ .

In the second case, we shall assume the following

H2:  $\mu$  has a density with respect to the Lebesgue measure,  
the  $\eta_k(\cdot, \theta_k)$  are continuously differentiable in  $x \in \text{int}(\mathcal{X})$ .

We assume that the models are distinguishable in the following sense:

$$\mu \{x \in \mathcal{X} : \eta_k(x, \theta_k) = \eta_j(x, \theta_j) \text{ for some } j \neq k\} = 0. \quad (2.1)$$

**Regret.** If the parameters  $\theta_k$  in each model  $\eta_k$  were known, we could use an oracle rule and allocate a subject with covariates  $X$  to treatment  $k^*$  such that

$$\eta_*(X) = \eta_{k^*}(X) = \max_{k=1, \dots, K} \eta_k(X, \theta_k).$$

However, for unknown  $\theta_k$  allocation to the best treatment cannot be guaranteed and we define the (cumulative) *regret after  $n$  allocations* as

$$R_n(\theta) = \frac{1}{n} \sum_{i=1}^n [\eta_*(X_i) - \eta_{k_i}(X_i, \theta_{k_i})],$$

when the  $i$ -th subject with covariates  $X_i$  has been allocated to treatment  $k_i$ , for all  $i = 1, \dots, n$ . This can also be written as

$$R_n(\theta) = \mathbb{P}_{\mu_n} \eta_* - \sum_{k=1}^K \mathbb{P}_{\xi_{n,k}} \eta_k(\cdot, \theta_k), \quad (2.2)$$

where, for a measure  $\nu$  on  $\mathcal{X}$  and a  $\nu$ -measurable function  $f : \mathcal{X} \rightarrow \mathbb{R}$  we denote  $\mathbb{P}_\nu f = \int_{\mathcal{X}} f(x) d\nu(x)$ . Here  $\mu_n$  is the empirical measure of the  $X_i$  and

$$\xi_{n,k} = \frac{N_{n,k}}{n} \mu_{n,k}, \quad k = 1, \dots, K, \quad (2.3)$$

with  $\mu_{n,k}$  the empirical measure of the  $X_i$  that have been allocated to treatment  $k$  in the first  $n$  assignments, and  $N_{n,k}$  their number. Other forms of regret will be suggested in Section 5.

**Information.** Let  $\theta \in \mathbb{R}^p$  denote the vector of all parameters in the  $K$  models  $\eta_k$ , with  $p < \sum_{k=1}^K \dim(\theta_k)$  when the models have some parameters in common. Information will be related to the precision of the Maximum Likelihood (ML) estimation of  $\theta$ , measured by the (inverse of the) normalized Fisher information matrix. Denote by  $\mathfrak{M}_k(x; \theta_k)$  the elementary information matrix corresponding to the observation of the response  $Y_t | X = x, t = k$ , with expectation  $\eta_k(x, \theta_k)$ , see (1.1);  $\mathfrak{M}_k(x; \theta_k)$  is  $p \times p$ , but its  $j$ -th row and column are formed of zeros when  $\theta_k$  does not contain the  $j$ -th component of  $\theta$ . For example, in Bernoulli trials with a single response  $Y_t \in \{0, 1\}$ , we have

$$\mathfrak{M}_k(x; \theta_k) = \frac{\partial \eta_k(x, \theta_k)}{\partial \theta} \frac{\partial \eta_k(x, \theta_k)}{\partial \theta^\top} \frac{1}{\eta_k(x, \theta_k)[1 - \eta_k(x, \theta_k)]}. \quad (2.4)$$

In case of H1b, we assume the following in complement of H2:

H2': all components of  $\mathfrak{M}_k(x; \theta_k)$  are continuously differentiable in  $x \in \text{int}(\mathcal{X})$ .

With notations similar to the regret calculation, we can compute the normalized information matrix after  $n$  allocations as

$$\mathbf{M}_n(\theta) = \frac{1}{n} \sum_{i=1}^n \mathfrak{M}_{k_i}(x; \theta_{k_i}) = \sum_{k=1}^K \mathbb{P}_{\xi_{n,k}} \mathfrak{M}_k(\cdot; \theta_k). \quad (2.5)$$

We shall measure the *information content* of the trial by  $\Psi(\mathbf{M}_n)$ , with  $\Psi(\cdot)$  a design criterion defined on the set of symmetric non-negative definite  $p \times p$  matrices. We suppose that  $\Psi(\cdot)$  is concave, monotonic for Loewner ordering, twice continuously differentiable and strictly concave on the set  $\mathbb{M}^+$  of symmetric positive-definite matrices. Typical examples are

$$\Psi_q(\mathbf{M}) = \begin{cases} -\text{trace}(\mathbf{M}^{-q}) & \text{for } q > 0, \\ \log \det(\mathbf{M}) & \text{for } q = 0 \text{ (D-optimality)}, \end{cases} \quad (2.6)$$

with  $A$ -optimal design when  $q = 1$ , see, e.g., Pukelsheim (1993, Chap. 5).

**Limiting allocations and design measures.** In the usual context of optimal design theory, the consideration of an asymptotic framework where design measures are substituted for exact designs of given size  $n$  much facilitates the construction of optimal designs. It is also the case here, with a further justification by the presence of the probability measure  $\mu$  at the core of the problem.

The measures  $\xi_{n,k}$  defined by (2.3) satisfy  $\sum_{k=1}^K \xi_{n,k} = \mu_n$ , with  $\mu_n \rightarrow \mu$  as  $n \rightarrow \infty$  since the  $X_i$  are i.i.d. with  $\mu$ . We shall thus consider design measures  $\underline{\xi} = (\xi_1, \dots, \xi_K)$  that form a decomposition of  $\mu$  into  $\mu = \sum_{k=1}^K \xi_k$ , where  $\xi_k$  will define the target limiting allocation in a sequential allocation rule and corresponds to the fraction of  $\mu$  devoted to treatment  $k$ . Note that the  $\xi_k$  are not probability measures. We shall denote by  $\Xi(\mu)$  the set of such  $\underline{\xi}$ ,

$$\Xi(\mu) = \{ \underline{\xi} = (\xi_1, \dots, \xi_K) \in (\mathcal{M}_{\mathcal{X}})^K : \sum_{k=1}^K \xi_k = \mu \}, \quad (2.7)$$

where  $\mathcal{M}_{\mathcal{X}}$  is the set of non-negative measures on  $\mathcal{X}$  absolutely continuous with respect to  $\mu$ . Notice that  $\Xi(\mu)$  is convex.

**Combining information and regret.** The regret  $R(\underline{\xi}; \theta)$  for a  $\underline{\xi} \in \Xi(\mu)$  can be written as

$$R(\underline{\xi}; \theta) = P_{\mu} \eta_* - \sum_{k=1}^K P_{\xi_k} \eta_k(\cdot, \theta_k), \quad (2.8)$$

see (2.2). The associated information is  $\psi(\underline{\xi}; \theta) = \Psi[\mathbf{M}(\underline{\xi}; \theta)]$ , with

$$\mathbf{M}(\underline{\xi}; \theta) = \sum_{k=1}^K P_{\xi_k} \mathfrak{M}_k(\cdot; \theta_k),$$

see (2.5). We suppose that  $\mathbf{M}(\underline{\xi}; \theta)$  is positive definite when all  $\xi_k = \mu/K$  in  $\underline{\xi}$ .

We shall consider design problems that correspond to the maximization of compromise criteria of the form  $J^{(\alpha)}(\underline{\xi}; \theta) = (1 - \alpha) \psi(\underline{\xi}; \theta) - \alpha R(\underline{\xi}; \theta)$  for some  $\alpha \in [0, 1]$ . In  $J^{(\alpha)}(\underline{\xi}; \theta)$  the information content of the trial, measured by  $\psi(\underline{\xi}; \theta)$ , is balanced by the regret  $R(\underline{\xi}; \theta)$  which corresponds to an ethical cost, see Pronzato (2010) for a similar approach in another context. Due to the equivalence between constrained and compound optimal designs, see Cook and Wong (1994); Pronzato (2010) and Fedorov and Leonov (2014, Chap. 4), this is equivalent to maximizing  $\psi(\underline{\xi}; \theta)$  under a constraint of the form  $R(\underline{\xi}; \theta) \leq \tau$  for some constant  $\tau$ . Indeed, the Lagrangian for this constrained problem can be written as  $\mathcal{L}(\underline{\xi}, C) = \psi(\underline{\xi}; \theta) - C [R(\underline{\xi}; \theta) - \tau]$ , with  $C \geq 0$  the Lagrange parameter; maximizing  $\psi(\underline{\xi}; \theta)$  with  $R(\underline{\xi}; \theta) \leq \tau$  is equivalent to maximizing  $\mathcal{L}(\underline{\xi}, C)$  for some  $C = C(\tau)$  and is equivalent to maximizing  $J^{(\alpha)}(\underline{\xi}; \theta)$  for  $\alpha = C/(1 + C)$ . Due to the concavity of  $\Psi(\cdot)$ , any non-dominated solution  $\underline{\xi}^*$  (that is, such that there does not exist a  $\underline{\xi}' \in \Xi(\mu)$  satisfying  $\psi(\underline{\xi}'; \theta) \geq \psi(\underline{\xi}^*; \theta)$  and  $R(\underline{\xi}'; \theta) \leq R(\underline{\xi}^*; \theta)$  with one of these inequalities being strict) maximizes  $J^{(\alpha)}(\underline{\xi}; \theta)$  for some  $\alpha \in [0, 1]$ . Figure 3—right in Example 3 will provide an illustration. Similarly, maximizing  $J^{(\alpha)}(\underline{\xi}; \theta)$  is equivalent to minimizing  $R(\underline{\xi}; \theta)$  under the constraint  $\psi(\underline{\xi}; \theta) \geq \tau'$  for some constant  $\tau'$ . Baldi Antognini and Giovagnoli (2010) give a justification for using compromise design in the context of clinical trials, but they do not take covariates into account. Balancing efficiency and ethics is also considered in (Hu et al., 2015), but no explicit optimal design

is used as a target for sequential allocation (see in particular Example 3). The achievement of a suitable balance between treatments is often explored through the introduction of randomization in the definition of the allocation rule, see, e.g., Ball et al. (1993); Atkinson and Biswas (2005); Hu and Hu (2012). The duality information/regret is then less clearly accounted for than in the maximization of  $J^{(\alpha)}(\underline{\xi}; \theta)$ . Randomization will be considered in Section 3 through the issue of selection bias.

In what follows, we omit the dependence in  $\theta$  of information and regret when it does not impede readability. Due to the expression (2.8) of the regret, maximizing  $J^{(\alpha)}(\underline{\xi})$  is equivalent to maximizing

$$H^{(\alpha)}(\underline{\xi}) = (1 - \alpha) \psi(\underline{\xi}) + \alpha \phi(\underline{\xi}) \quad (2.9)$$

where  $\phi(\underline{\xi})$  is a cumulative reward (to be maximized),  $\phi(\underline{\xi}) = \sum_{k=1}^K P_{\xi_k} \eta_k(\cdot)$ . Notice that when  $\Psi(\mathbf{M}) = \log \det(\mathbf{M})$  ( $D$ -optimality), optimal designs for  $H^{(\alpha)}(\cdot)$  are invariant by reparameterization of the models  $\eta_k$  for any  $\alpha \in [0, 1]$ .

## 2.2. An Equivalence Theorem for compromise optimal designs

For any concave functional  $h(u)$  defined for all  $u$  in some linear space  $\mathcal{S}$  and for any  $u$  and  $v$  in  $\mathcal{S}$ , we shall denote  $G_h(u; v) = \lim_{\gamma \rightarrow 0^+} [h(u + \gamma v) - h(u)]/\gamma$  and  $F_h(u; v) = G_h(u; v - u)$ . Due to the concavity of  $h(\cdot)$  these directional derivatives exist in  $\mathbb{R} \cup \{+\infty\}$  for any  $u$  such that  $h(u) > -\infty$ , see, e.g., Pshenichnyi (1971, p. 38), Pronzato and Pázman (2013, Lemma 5.16). For any  $\underline{\xi} \in (\mathcal{M}_{\mathcal{X}})^K$  (not necessarily in  $\Xi(\mu)$  given by (2.7)) such that  $\Psi(\cdot)$  is differentiable at  $\mathbf{M}(\underline{\xi})$ , we denote  $\nabla \psi(\underline{\xi}) = \nabla \Psi[\mathbf{M}(\underline{\xi})]$ , where  $\nabla \Psi(\mathbf{M}) = \partial \Psi(\mathbf{M})/\partial \mathbf{M}$ . For any  $\underline{\xi}$  and  $\underline{\nu}$  in  $(\mathcal{M}_{\mathcal{X}})^K$  such that  $\mathbf{M}(\underline{\xi}) \in \mathbb{M}^+$ , we then have

$$G_{H^{(\alpha)}}(\underline{\xi}; \underline{\nu}) = (1 - \alpha) \text{trace}[\nabla \psi(\underline{\xi}) \mathbf{M}(\underline{\nu})] + \alpha \phi(\underline{\nu}) = \sum_{k=1}^K G_k^{(\alpha)}(\underline{\xi}; \nu_k),$$

where  $G_k^{(\alpha)}(\underline{\xi}; \nu_k) = (1 - \alpha) \text{trace}[\nabla \psi(\underline{\xi}) \mathbf{M}_k(\nu_k)] + \alpha \phi_k(\nu_k)$ , with  $\mathbf{M}_k(\nu_k) = P_{\nu_k} \mathfrak{M}_k(\cdot)$  and  $\phi_k(\nu_k) = P_{\nu_k} \eta_k(\cdot)$ ,  $k = 1, \dots, K$ . When  $\psi(\cdot)$  is differentiable at  $\underline{\xi}$  we thus obtain

$$G_k^{(\alpha)}(\underline{\xi}; \nu_k) = P_{\nu_k} G_k^{(\alpha)}(\underline{\xi}, \cdot) = \int_{\mathcal{X}} G_k^{(\alpha)}(\underline{\xi}, x) d\nu_k(x),$$

where we have denoted

$$G_k^{(\alpha)}(\underline{\xi}, x) = G_k^{(\alpha)}(\underline{\xi}; \delta_x) = (1 - \alpha) \text{trace}[\nabla \psi(\underline{\xi}) \mathfrak{M}_k(x)] + \alpha \eta_k(x), \quad (2.10)$$

with  $\delta_x$  the delta measure at  $x$ . As usual in design theory, the convexity of  $\Xi(\mu)$  given by (2.7) and the concavity of  $H^{(\alpha)}(\cdot)$  given by (2.9) yield an Equivalence Theorem, which states that  $\underline{\xi}^* \in \Xi(\mu)$  maximizes  $H^{(\alpha)}(\cdot)$  if and only if  $F_{H^{(\alpha)}}(\underline{\xi}^*; \underline{\nu}) \leq 0$  for any  $\underline{\nu} \in \Xi(\mu)$ , where

$$F_{H^{(\alpha)}}(\underline{\xi}; \underline{\nu}) = (1 - \alpha) \text{trace}[\nabla \psi(\underline{\xi}) \{\mathbf{M}(\underline{\nu}) - \mathbf{M}(\underline{\xi})\}] + \alpha [\phi(\underline{\nu}) - \phi(\underline{\xi})].$$

This necessary and sufficient condition for optimality can be restated as follows.

**Theorem 2.1** *Suppose that  $H^{(\alpha)}(\cdot)$  is differentiable at  $\underline{\xi}^{*(\alpha)} = (\xi_1^*, \dots, \xi_K^*) \in \Xi(\mu)$ . The following statements are equivalent:*

- (i)  $\underline{\xi}^* = \underline{\xi}^{*(\alpha)}$  is optimal, i.e.,  $\underline{\xi}^*$  maximizes  $H^{(\alpha)}(\underline{\xi})$  with respect to  $\underline{\xi} \in \Xi(\mu)$ ;
- (ii) for all  $i = 1, \dots, K$ ,  $G_i^{(\alpha)}(\underline{\xi}^*, x) \geq \max_{j \neq i} G_j^{(\alpha)}(\underline{\xi}^*, x)$   $\xi_i^*$ -a.e.;
- (iii)  $\mathcal{X}$  can be partitioned into  $n_K$  subsets  $\mathcal{X}_{\mathcal{J}_t} = \mathcal{X}_{\mathcal{J}_t}^{(\alpha)}$ ,  $t = 1, \dots, n_K \leq 2^K - 1$ , with index sets  $\mathcal{J}_t$  that are subsets of  $\{1, \dots, K\}$ , and such that, for all  $t = 1, \dots, n_K$ ,
  - (a)  $\sum_{i \in \mathcal{J}_t} \xi_i^* = \mu$  on  $\mathcal{X}_{\mathcal{J}_t}$ ,
  - (b)  $G_i^{(\alpha)}(\underline{\xi}^*, x) = G_j^{(\alpha)}(\underline{\xi}^*, x)$  on  $\mathcal{X}_{\mathcal{J}_t}$  for all  $i, j \in \mathcal{J}_t$ ,
  - (c)  $G_i^{(\alpha)}(\underline{\xi}^*, x) > G_j^{(\alpha)}(\underline{\xi}^*, x)$  for  $x \in \mathcal{X}_{\mathcal{J}_t}$ ,  $i \in \mathcal{J}_t$  and  $j \notin \mathcal{J}_t$ .

The proof of Theorem 2.1 is given in Appendix. The theorem takes a simpler form when  $K = 2$ : the function  $x \rightarrow \Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) = G_1^{(\alpha)}(\underline{\xi}^*, x) - G_2^{(\alpha)}(\underline{\xi}^*, x)$  then defines a partition of  $\mathcal{X}$  in  $m$  sets,  $m \leq 3$ , as expressed in the following corollary.

**Corollary 2.1** *Suppose that  $K = 2$  and  $H^{(\alpha)}(\cdot)$  is differentiable at  $\underline{\xi}^{*(\alpha)} = (\xi_1^*, \xi_2^*) \in \Xi(\mu)$ . The following statements are equivalent:*

- (i)  $\underline{\xi}^* = \underline{\xi}^{*(\alpha)}$  is optimal, i.e., it maximizes  $H^{(\alpha)}(\underline{\xi})$  with respect to  $\underline{\xi} \in \Xi(\mu)$ ;
- (ii)  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) \geq 0$   $\xi_1^*$ -a.e. and  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) \leq 0$   $\xi_2^*$ -a.e.;
- (iii) there exist two subsets  $\mathcal{X}_1 = \mathcal{X}_1^{(\alpha)}$  and  $\mathcal{X}_2 = \mathcal{X}_2^{(\alpha)}$  of  $\mathcal{X}$  such that
  - (a)  $\xi_1^* = \mu$  on  $\mathcal{X}_1$  and  $\xi_2^* = \mu$  on  $\mathcal{X}_2$ ,
  - (b)  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) = 0$  on  $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$ ,
  - (c)  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) > 0$  for  $x \in \mathcal{X}_1$  and  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) < 0$  for  $x \in \mathcal{X}_2$ .

Notice that  $\underline{\xi}^*$ , the  $K$  functions  $x \rightarrow G_k^{(\alpha)}(\underline{\xi}, x)$  and the sets  $\mathcal{X}_i$  depend on  $\theta$ , see (2.10). We shall use the notations  $\underline{\xi}_{\theta}^*$ ,  $G_k^{(\alpha)}(\underline{\xi}, x; \theta)$ ,  $\mathcal{X}_i(\theta)$  to emphasize this dependence when necessary (Section 5).

### 2.3. Some properties of compromise optimal designs

The values of  $H^{(\alpha)}(\underline{\xi}^{*(\alpha)})$ ,  $(1 - \alpha)\psi(\underline{\xi}^{*(\alpha)})$  and  $\alpha\phi(\underline{\xi}^{*(\alpha)})$  are always uniquely defined for any  $\alpha \in [0, 1]$  (which implies that  $\psi(\underline{\xi}^{*(\alpha)})$  and  $\phi(\underline{\xi}^{*(\alpha)})$  are uniquely defined for  $\alpha$  respectively in  $[0, 1)$  and  $(0, 1]$ ). Moreover, one can show that  $H^{(\alpha)}(\underline{\xi}^{*(\alpha)})$ ,  $(1 - \alpha)\psi(\underline{\xi}^{*(\alpha)})$  and  $\alpha\phi(\underline{\xi}^{*(\alpha)})$  are continuous functions of  $\alpha$  in  $[0, 1]$ , that  $\psi(\underline{\xi}^{*(\alpha)})$  is non-increasing for  $\alpha \in [0, 1)$  and  $\phi(\underline{\xi}^{*(\alpha)})$  is non-decreasing on  $(0, 1]$ , and that  $H^{(\alpha)}(\underline{\xi}^{*(\alpha)})$  is convex and continuously differentiable with respect to  $\alpha$  in  $(0, 1)$ , with  $dH^{(\alpha)}(\underline{\xi}^{*(\alpha)})/d\alpha = \phi(\underline{\xi}^{*(\alpha)}) - \psi(\underline{\xi}^{*(\alpha)})$ .

### 2.3.1. The special case $\alpha = 1$

This case corresponds to the usual framework in bandit theory, with abundant results on the construction of strategies minimizing the expected regret, see, e.g., Lai and Robbins (1985); Goldenshluger and Zeevi (2011, 2013). Theorem 2.1 applies when  $\alpha = 1$  since  $H^{(1)}(\underline{\xi}) = \sum_{k=1}^K P_{\xi_k} \eta_k(\cdot)$  is linear in the  $\xi_k$ , and the sets  $\mathcal{X}_j$  are uniquely defined. For instance, when  $K = 2$ , (2.1) implies that the optimal design is given by  $\underline{\xi}^* = \underline{\xi}^{*(1)} = (\xi_1^*, \xi_2^*)$  such that  $\xi_1^* = \mu$  on  $\mathcal{X}_1^{(1)} = \{x \in \mathcal{X} : \eta_1(x) > \eta_2(x)\}$  and  $\xi_2^* = \mu$  on  $\mathcal{X}_2^{(1)} = \{x \in \mathcal{X} : \eta_1(x) < \eta_2(x)\}$ . Therefore, when the models are such that  $\eta_1(x) > \eta_2(x)$  for all  $x \in \mathcal{X}$ ,  $\mathcal{X}_1^{(1)} = \mathcal{X}$  and  $\underline{\xi}^*$  does not allow estimation of  $\theta_2$ . In a sequential response-adaptive situation where assignment decisions are based on estimated values of the model parameters, it means that a deterministic decision rule using  $\alpha = 1$  (a method sometimes called “best intention design”) may fail to ensure the consistent estimation of  $\theta$ ; moreover, allocation to the poorest treatment for all  $n$  large enough may happen with positive probability.

When  $\mathbf{M}(\underline{\xi}^{*(1)})$  is nonsingular, the design  $\underline{\xi}^{*(1)}$  may be optimal for all  $H^{(\alpha)}(\cdot)$  with  $\alpha$  in some interval  $[\underline{\alpha}, 1]$ , see Example 3. Note that

$$\max_{\underline{\xi} \in \Xi(\mu)} \phi(\underline{\xi}) = \phi(\underline{\xi}^{*(1)}) \text{ and } R(\underline{\xi}) = \phi(\underline{\xi}^{*(1)}) - \phi(\underline{\xi}) \text{ for any } \underline{\xi} \in \Xi(\mu). \quad (2.11)$$

### 2.3.2. Uniqueness of $\mathbf{M}(\underline{\xi}^{*(\alpha)})$ for $\alpha < 1$

The information criteria (2.6) are such that  $\Psi(\mathbf{M}) = -\infty$  for any singular  $\mathbf{M}$  and is finite otherwise. Then, for any  $\alpha \in [0, 1)$ , a design  $\underline{\xi}^{*(\alpha)}$  optimal for  $H^{(\alpha)}(\cdot)$  defined by (2.9) is such that  $\mathbf{M}(\underline{\xi}^{*(\alpha)}) \in \mathbb{M}^+$ ;  $H^{(\alpha)}(\cdot)$  is thus differentiable at  $\underline{\xi}^{*(\alpha)}$  and Theorem 2.1 applies. This is also the case for the positively homogeneous versions  $\Psi(\mathbf{M}) = \det^{1/p}(\mathbf{M})$  and  $\Psi(\mathbf{M}) = (\text{trace}[\mathbf{M}^{-q}]/p)^{-1/q}$  ( $q > 0$ ), which have continuous extensions  $\Psi(\mathbf{M}) = 0$  at singular  $\mathbf{M}$ , see Pukelsheim (1993, Chap. 5). Indeed, the property that the directional derivative  $F_{\Psi}(\mathbf{M}, \mathbf{M}')$  equals  $+\infty$  when  $\mathbf{M}$  is singular and  $\mathbf{M}'$  has full rank ensures that  $\mathbf{M}(\underline{\xi}^{*(\alpha)})$  is nonsingular when  $\alpha < 1$ .

The strict concavity of  $\Psi(\cdot)$  implies the uniqueness of  $\mathbf{M}(\underline{\xi}^{*(\alpha)}) \in \mathbb{M}^+$ , and therefore of the functions  $x \rightarrow G_k^{(\alpha)}(\underline{\xi}^{*(\alpha)}, x)$  and sets  $\mathcal{X}_{\mathcal{J}_t}^{(\alpha)}$  in Theorem 2.1-*iii*).

### 2.3.3. Uniqueness of $\underline{\xi}^{*(\alpha)}$

The optimal design  $\underline{\xi}^{*(\alpha)} = (\xi_1^*, \dots, \xi_K^*)$  that maximizes  $H^{(\alpha)}(\underline{\xi})$  is uniquely defined (up to sets of zero measure) when

$$\mu\{x \in \mathcal{X} : G_i^{(\alpha)}(\underline{\xi}^*, x) = G_j^{(\alpha)}(\underline{\xi}^*, x)\} = 0 \text{ for all } i \neq j \quad (2.12)$$

(which corresponds to  $\mu(\mathcal{X}_1 \cup \mathcal{X}_2) = 1$  when  $K = 2$ , see Corollary 2.1).

Condition (2.12) is often satisfied when  $\alpha \in [0, 1)$  and  $\mu$  has a density with respect to the Lebesgue measure, but the reverse situation cannot be considered as exceptional, as Example 1 will illustrate. A simple generic case where the condition fails is when  $\alpha = 0$ ,  $K = 2$ ,  $\eta_1, \eta_2$  have no parameters in common, but the numerical values of  $\theta_1$  and  $\theta_2$  are such that, for any  $\xi$ ,  $\mathbf{M}_1(\xi) = c\mathbf{M}_2(\xi)$  when we write  $\mathbf{M}(\xi)$  as

$$\mathbf{M}(\xi) = \begin{pmatrix} \mathbf{M}_1(\xi_1) & \mathbf{O} \\ \mathbf{O}^\top & \mathbf{M}_2(\xi_2) \end{pmatrix},$$

with  $c$  some positive constant. Take for instance  $\Psi(\mathbf{M}) = \log \det(\mathbf{M})$ . Then, for any  $\xi \in \Xi(\mu)$  we have  $H^{(0)}(\xi) = \log \det[\mathbf{M}_1(\xi_1)] + \log \det[\mathbf{M}_2(\xi_2)] + p_1 \log(c)$ , with  $p_1$  the number of parameters in  $\eta_1$  (and  $\eta_2$ ). Consider  $\xi_{\underline{\mu}} = (\mu/2, \mu/2) = ((\xi_1 + \xi_2)/2, (\xi_1 + \xi_2)/2) \in \Xi(\mu)$ ; it satisfies  $H^{(0)}(\xi_{\underline{\mu}}) = 2 \log \det\{\mathbf{M}_1[(\xi_1 + \xi_2)/2]\} + p_1 \log(c)$ . The concavity of  $\log \det(\cdot)$  implies that  $H^{(0)}(\xi_{\underline{\mu}}) \geq H^{(0)}(\xi)$ , and  $\xi_{\underline{\mu}}$  is thus optimal, with  $\mathcal{X}_1 = \mathcal{X}_2 = \emptyset$  in Corollary 2.1-(iii). Moreover, any optimal design  $\xi^* = (\xi_1^*, \xi_2^*)$  is such that  $\mathbf{M}_1(\xi_1^*) = \mathbf{M}_1(\xi_2^*)$ , and the designs  $(\xi_2^*, \xi_1^*)$  and  $(1-\gamma)\xi^* + \gamma\xi_{\underline{\mu}}$  are also optimal for all  $\gamma \in [0, 1]$ . Since  $\alpha = 0$ , these optimal designs may have different regret values, see Example 3.

When (2.12) is satisfied and  $\xi^{*(\alpha)}$  is uniquely defined, the sequential allocation rules presented in Section 4 are such that the empirical measures  $\xi_{n,k}$  defined by (2.3) converge a.s. to the  $\xi_k^{*(\alpha)}$  (weak convergence), and the allocation proportions  $\xi_{n,k}(\mathcal{X})$  converge a.s. to their optimal counterparts  $\xi_k^{*(\alpha)}(\mathcal{X})$ . Note that when  $\alpha < 1$  this convergence of allocation proportions can always be ensured by including an intercept  $\theta_{0k}$  in each of the  $k$  models  $\eta_k$  due to the uniqueness of  $\mathbf{M}(\xi^{*(\alpha)})$ , see Section 2.3.2.

#### 2.4. Bounds on optimal regret and information

**Upper bounds on optimal regret.** Let  $\xi^{*(\alpha)} \in \Xi(\mu)$  be an optimal design that maximizes  $H^{(\alpha)}(\xi)$  for a given  $\alpha \in (0, 1]$ . Take any  $k \in \{1, \dots, K\}$  and consider a set  $\mathcal{X}_{\mathcal{J}_t}$  in Theorem 2.1-(iii) that contains  $k$ , the measure  $\xi_k$  being positive on such sets  $\mathcal{X}_{\mathcal{J}_t}$  only. For any  $x \in \mathcal{X}_{\mathcal{J}_t}$  and any  $j \in \{1, \dots, K\}$  we have

$$\eta_j(x) - \eta_k(x) \leq \frac{1-\alpha}{\alpha} \text{trace} \left\{ \nabla \psi(\xi^{*(\alpha)}) [\mathfrak{M}_k(x) - \mathfrak{M}_j(x)] \right\},$$

see (2.10). The monotonicity of  $\Psi(\cdot)$  for Loewner ordering implies that  $\nabla \psi(\xi^{*(\alpha)})$  is non-negative definite, therefore

$$\eta_*(x) - \eta_k(x) = \max_j [\eta_j(x) - \eta_k(x)] \leq \frac{1-\alpha}{\alpha} \text{trace} \left[ \nabla \psi(\xi^{*(\alpha)}) \mathfrak{M}_k(x) \right].$$

Repeating the same operation for all  $k$ , and integrating over all  $\xi_k^{*(\alpha)}$ , we obtain

$$R(\xi^{*(\alpha)}) \leq \frac{1-\alpha}{\alpha} \text{trace} \left[ \nabla \psi(\xi^{*(\alpha)}) \mathbf{M}(\xi^{*(\alpha)}) \right]. \quad (2.13)$$

In the special case  $\Psi(\mathbf{M}) = \log \det(\mathbf{M})$  the bound does not depend on  $\underline{\xi}^{*(\alpha)}$ : indeed, we then have  $\nabla \psi(\underline{\xi}) = \mathbf{M}^{-1}(\underline{\xi})$ , which gives  $R(\underline{\xi}^{*(\alpha)}) \leq p(1 - \alpha)/\alpha$ .

Another upper bound on  $R(\underline{\xi}^{*(\alpha)})$  can be obtained by exploiting the concavity of  $\psi(\cdot)$  and comparing  $\underline{\xi}^{*(\alpha)}$  with  $\underline{\xi}^{*(1)}$ . For any  $\alpha \in [0, 1]$ , we have  $H^{(\alpha)}(\underline{\xi}^{*(\alpha)}) \geq H^{(\alpha)}(\underline{\xi}^{*(1)})$ , so that, using (2.11),

$$\begin{aligned} R(\underline{\xi}^{*(\alpha)}) &= \phi(\underline{\xi}^{*(1)}) - \phi(\underline{\xi}^{*(\alpha)}) \leq \frac{1 - \alpha}{\alpha} [\psi(\underline{\xi}^{*(\alpha)}) - \psi(\underline{\xi}^{*(1)})] \\ &\leq \frac{1 - \alpha}{\alpha} \text{trace} \left\{ \nabla \psi(\underline{\xi}^{*(1)}) [\mathbf{M}(\underline{\xi}^{*(\alpha)}) - \mathbf{M}(\underline{\xi}^{*(1)})] \right\}. \end{aligned} \quad (2.14)$$

**A lower bound on optimal information.** By definition, an optimal design  $\underline{\xi}^{*(\alpha)}$  maximizing  $H^{(\alpha)}(\underline{\xi})$ ,  $\alpha \in [0, 1)$ , satisfies  $\psi(\underline{\xi}^{*(\alpha)}) \geq \psi(\underline{\xi}) + \alpha[\phi(\underline{\xi}) - \phi(\underline{\xi}^{*(\alpha)})]/(1 - \alpha)$  for all  $\underline{\xi} \in \Xi(\mu)$ . Therefore,  $\psi(\underline{\xi}^{*(\alpha)}) \geq \psi(\underline{\xi}) + \alpha[\phi(\underline{\xi}) - \max_{\underline{\nu} \in \Xi(\mu)} \phi(\underline{\nu})]/(1 - \alpha)$ . Using (2.11), we obtain

$$\psi(\underline{\xi}^{*(\alpha)}) \geq \psi(\underline{\xi}) - \frac{\alpha}{1 - \alpha} R(\underline{\xi}), \quad \forall \underline{\xi} \in \Xi(\mu). \quad (2.15)$$

## 2.5. Examples

### 2.5.1. Example 1

Consider a linear response bandit problem, with  $K = 2$ , and  $\eta_k(x, \theta_k) = a_0 + b_k x$ ,  $k = 1, 2$ . The two models have the parameter  $a_0$  in common, and  $\theta = (a_0, b_1, b_2)^\top$ . We take  $\Psi(\mathbf{M}) = \log \det(\mathbf{M})$  and suppose that  $\mu$  is uniform on  $\mathcal{X} = [-1, 1]$  and that  $Y_t | X = x, t = k$  has mean  $\eta_k(x, \theta_k)$  and variance 1. We thus have, for any  $x \in \mathcal{X}$  and any  $\underline{\xi} = (\xi_1, \xi_2) \in \Xi(\mu)$ ,

$$\mathfrak{M}_1(x) = \begin{pmatrix} 1 & x & 0 \\ x & x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathfrak{M}_2(x) = \begin{pmatrix} 1 & 0 & x \\ 0 & 0 & 0 \\ x & 0 & x^2 \end{pmatrix}, \quad \mathbf{M}(\underline{\xi}) = \begin{pmatrix} 1 & m_1 & m_2 \\ m_1 & s_1^2 & 0 \\ m_2 & 0 & s_2^2 \end{pmatrix},$$

where  $m_k = \int_{\mathcal{X}} x \, d\xi_k(x)$ ,  $s_k^2 = \int_{\mathcal{X}} x^2 \, d\xi_k(x)$  (with  $m_1 + m_2 = 0$  and  $s_1^2 + s_2^2 = \int_{\mathcal{X}} x^2 \, d\mu(x) = 1/3$ ). This example illustrates the fact that we can have  $\mathcal{X}_1^{(\alpha)} = \mathcal{X}_2^{(\alpha)} = \emptyset$  in Corollary 2.1 for all  $\alpha$  in some interval also in the case where  $\mu$  has a density with respect to the Lebesgue measure. Without any loss of generality, we suppose that  $a_0 = 0$  and  $b_2 > b_1$ .

Direct calculation shows that the optimal design for all  $\alpha \geq \underline{\alpha} = 24/(24 + b_2 - b_1)$  is  $\underline{\xi}^{*(1)}$  such that  $\mathcal{X}_1^{(\alpha)} = [-1, 0)$  and  $\mathcal{X}_2^{(\alpha)} = (0, 1]$ .

On the other hand,  $G_1^{(\alpha)}(\underline{\xi}, x)$  and  $G_2^{(\alpha)}(\underline{\xi}, x)$  are polynomials of degree 2 in  $x$  for any  $\underline{\xi} \in \Xi(\mu)$ , and when  $\alpha \leq \underline{\alpha}$  their difference can be made identically zero for all  $x \in \mathcal{X}$  by choosing a  $\underline{\xi} \in \Xi(\mu)$  with suitable values of  $m_1, m_2, s_1^2$  and  $s_2^2$ .

Then,  $\mathcal{X}_1^{(\alpha)} = \mathcal{X}_2^{(\alpha)} = \emptyset$  in Corollary 2.1. The conditions are  $s_1^2 = s_2^2 = 1/6$  and  $m_1 = -m_2 = m_1^{*(\alpha)}$  with

$$m_1^{*(\alpha)} = \frac{1}{\alpha(b_2 - b_1)} \left( 1 - \alpha - \left[ (1 - \alpha)^2 + \frac{\alpha^2(b_2 - b_1)^2}{12} \right]^{1/2} \right).$$

They are fulfilled for instance for  $\xi_1^{*(\alpha)} = \mu$  on  $[-1, -A(\alpha)] \cup [0, A(\alpha)]$  and  $\xi_2^{*(\alpha)} = \mu$  on the complement, with  $A(\alpha) = \sqrt{2}(m_1^{*(\alpha)} + 1/4)^{1/2}$ , which satisfies  $A(\underline{\alpha}) = 0$  and  $\lim_{\alpha \rightarrow 0} A(\alpha) = 1/\sqrt{2}$ . Also, elementary calculations show that the optimal solution whose (Shannon) entropy  $\sum_{i=1}^2 \int_{\mathcal{X}} -\log\{[d\xi_i^*/d\mu](x)\} d\xi_i^*(x)$  is maximum is given by  $[d\xi_1^*/d\mu](x) = (1/2)[1 + \exp(\lambda_1 x)]^{-1}$ , with  $\lambda_1 > 0$  such that  $m_1 = m_1^{*(\alpha)}$ . Note that  $\xi_1^{*(\alpha)}(\mathcal{X}) = \xi_2^{*(\alpha)}(\mathcal{X}) = 1/2$  for all  $\alpha$  for these two types of solutions, but other optimal designs  $\underline{\xi}^{*(\alpha)}$  exist such that  $\xi_1^{*(\alpha)}(\mathcal{X}) \neq \xi_2^{*(\alpha)}(\mathcal{X})$  for  $\alpha < \underline{\alpha}$ .

The information and regret of a design  $\underline{\xi}$  only depend on  $m_1$  and  $s_1^2$  and their optimal values are given by

$$\begin{aligned} \psi(\underline{\xi}^{*(\alpha)}) &= -2 \log(12), \quad R(\underline{\xi}^{*(\alpha)}) = 0 \quad \text{for } \alpha \geq \underline{\alpha} \\ \psi(\underline{\xi}^{*(\alpha)}) &= \log \left[ \frac{1}{36} - \frac{1}{3} (m_1^{*(\alpha)})^2 \right], \quad R(\underline{\xi}^{*(\alpha)}) = (b_2 - b_1) \left[ m_1^{*(\alpha)} + \frac{1}{4} \right] \quad \text{otherwise.} \end{aligned}$$

### 2.5.2. Example 2

A slightly simpler version of previous example yields a completely different solution. Take  $\eta_1(x, \theta_k) = a_0$ ,  $\eta_2(x, \theta_2) = a_0 + b_2 x$ , so that there are two parameters only,  $\theta = (a_0, b_2)^\top$ . We still take  $\mu$  uniform on  $\mathcal{X} = [-1, 1]$ ,  $\Psi(\mathbf{M}) = \log \det(\mathbf{M})$ , and suppose that  $b_2 > 0$ . The information matrix for a  $\underline{\xi} \in \Xi(\mu)$  is

$$\mathbf{M}(\underline{\xi}) = \begin{pmatrix} 1 & m_2 \\ m_2 & s_2^2 \end{pmatrix}$$

with  $m_2 = \int_{\mathcal{X}} x d\xi_2(x)$ ,  $s_2^2 = \int_{\mathcal{X}} x^2 d\xi_2(x)$ . The design  $\underline{\xi}^{*(1)}$  with  $\mathcal{X}_1^{(\alpha)} = [-1, 0]$  and  $\mathcal{X}_2^{(\alpha)} = (0, 1]$  is now optimal for  $\alpha \geq \underline{\alpha} = 72/(72 + 5b_2)$ . For  $\alpha < \underline{\alpha}$ ,  $G_2^{(\alpha)}(\underline{\xi}, x)$  is still a polynomial of degree 2 in  $x$  but  $G_1^{(\alpha)}(\underline{\xi}, x)$  is constant, and their difference cannot be made identically zero for  $x$  in an interval. The solution is thus much different from that in Example 1: the optimal design  $\underline{\xi}^{*(\alpha)}$  is uniquely defined for  $\alpha < \underline{\alpha}$ , it corresponds to  $\xi_1^{*(\alpha)} = \mu$  on  $\mathcal{X}_1^{(\alpha)} = (A(\alpha), 0)$  and  $\xi_2^{*(\alpha)} = \mu$  on  $\mathcal{X}_2^{(\alpha)} = [-1, A(\alpha)) \cup (0, 1]$ , where  $A(\alpha) \in [-1, 0]$  is solution of the fourth-degree equation

$$3\alpha b_2 A^4 - 8\alpha b_2 A^3 + 24(1 - \alpha)A^2 - 48(1 - \alpha)A - 16\alpha b_2 = 0,$$

with  $A(\underline{\alpha}) = -1$  and  $A(0) = 0$  (for  $\alpha = 0$ , pulling only the second arm permits to estimate both  $a_0$  and  $b_2$  while regret is ignored).

2.5.3. Example 3

We consider optimal allocation for two logistic regression models with  $\mathfrak{M}_k(x, \theta_k)$  given by (2.4),  $k = 1, 2$ , and  $\eta_1(x) = \eta_1(x, \theta_1) = 0.25 + 0.5 e^{z_1(x)}/(1 + e^{z_1(x)})$ ,  $\eta_2(x) = \eta_2(x, \theta_2) = 0.25 + 0.5/(1 + e^{z_2(x)})$ , where  $z_i(x) = b_i(x - a_i)$ ,  $\theta_i = (a_i, b_i)$ ,  $i = 1, 2$ .

We take  $\Psi(\mathbf{M}) = \log \det(\mathbf{M})$ , with nominal parameter values  $\theta_1 = \theta_2 = (1/2, 10)^\top$ , and  $\mu$  uniform on  $\mathcal{X} = [0, 1]$ . Figure 1 presents  $\eta_1(x)$  and  $\eta_2(x)$  as functions of  $x$ , with  $\eta_1(x) = \eta_2(x)$  for  $x = 1/2$ . The two responses are chosen totally different on purpose.

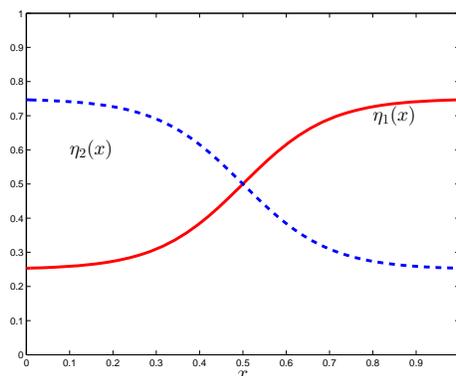


FIG 1. Example 3:  $\eta_1(x)$  (solid line) and  $\eta_2(x)$  (dashed line) as functions of  $x$ .

For any  $\alpha \in (0, 1]$  the optimal design  $\underline{\xi}^* = \underline{\xi}^{*(\alpha)}$  is uniquely defined, with  $\xi_1^* = \mu$  on  $\mathcal{X}_1^{(\alpha)} = (A(\alpha), B(\alpha)) \cup (C(\alpha), 1]$  and  $\xi_2^* = \mu$  on  $\mathcal{X}_2^{(\alpha)} = [0, A(\alpha)) \cup (B(\alpha), C(\alpha))$  for some  $A(\alpha) \leq B(\alpha) \leq C(\alpha)$  in  $(0, 1)$ . Figure 2-left illustrates the situation for  $\alpha = 0.7$  through a plot of  $G_1^{(\alpha)}(\underline{\xi}^*, x) - G_2^{(\alpha)}(\underline{\xi}^*, x)$  as a function of  $x$ , see Corollary 2.1-(ii). The optimal designs obtained for  $\alpha$  varying in  $(0, 1]$  are shown in Fig. 2-right. Here  $\eta_2(x) = 1 - \eta_1(x)$ ; the peculiar symmetry of the responses yields  $A(\alpha) + C(\alpha) = 1$  and  $B(\alpha) = 1/2$  for all  $\alpha$ .

The optimal design  $\underline{\xi}^{*(1)}$  for  $\alpha = 1$  corresponds to  $A(1) = C(1) = 1/2$  and  $\xi_1^{*(1)} = \mu$  on  $(1/2, 1]$ ,  $\xi_2^{*(1)} = \mu$  on  $[0, 1/2)$ . Numerical calculations show that it is optimal for all  $H^{(\alpha)}(\cdot)$  with  $\alpha \geq \underline{\alpha} \simeq 0.9949$ .

Figure 3-left presents the optimal regret  $R(\underline{\xi}^{*(\alpha)})$  (solid line) and the upper bound  $p(1 - \alpha)/\alpha$  (dashed line) as functions of  $\alpha$ ; the bottom part of the figure shows  $\psi(\underline{\xi}^{*(\alpha)})$  (solid line) and the lower bound (2.15) (dashed line) obtained for  $\underline{\xi} = \underline{\xi}_\mu = (\mu/2, \mu/2)$ . Since  $\eta_2(x) = 1 - \eta_1(x)$ , we are in the situation of Section 2.3.3: when  $\alpha = 0$ , any convex combination  $(1 - \gamma)\underline{\xi}^{*(0)} + \gamma\underline{\xi}_\mu$  with  $\gamma \in [0, 1]$  is optimal, where  $\underline{\xi}^{*(0)}$  denotes the particular solution given by  $\xi_1^{*(0)} = \mu$  on  $\mathcal{X}_1^{(0)} \simeq (0.3621, 0.5) \cup (0.6379, 1]$  and  $\xi_2^{*(0)} = \mu$  on  $\mathcal{X}_2^{(0)} \simeq [0, 0.3621) \cup$

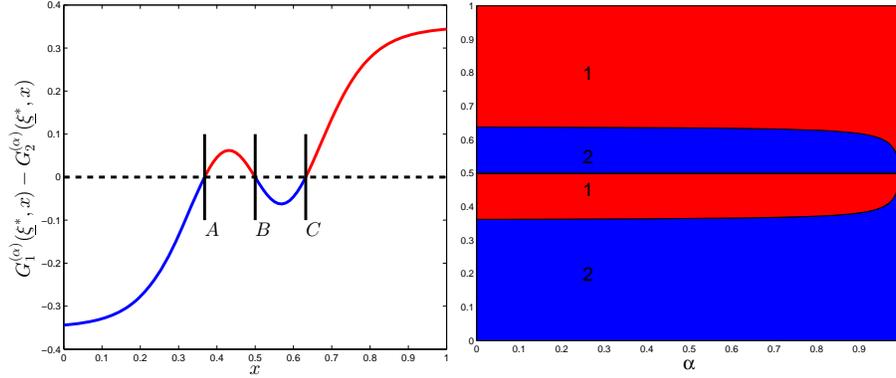


FIG 2. Example 3 — Left: illustration of the optimality of  $\underline{\xi}^*$  for  $\alpha = 0.7$ ; right: subsets  $\mathcal{X}_1^{(\alpha)}$  and  $\mathcal{X}_2^{(\alpha)}$  of Corollary 2.1-(iii) for  $\alpha \in (0, 1]$ .

(0.5, 0.6379). The regret at  $\alpha = 0$  can take any value between  $R(\underline{\xi}^{*(0)})$  and  $R(\underline{\xi}_\mu)$ .

Figure 3-right presents the information  $\psi(\underline{\xi}^{*(\alpha)})$  as a function of the regret  $R(\underline{\xi}^{*(\alpha)})$  for  $\alpha \in (0, 1]$ . Non-dominated solutions (see Section 2.1) correspond to the curve in solid-line, on which the solution for  $\alpha = 0.95$  is indicated by a circle. The slope of the tangent to the curve at this point (in dashed line) equals  $C = \alpha/(1 - \alpha) = 19$ , with  $C$  the Lagrange coefficient for the maximization of  $\Psi(\underline{\xi})$  under the constraint  $R(\underline{\xi}) \leq R(\underline{\xi}^{*(0.95)})$ .

The construction of optimal designs  $\underline{\xi}^{*(\alpha)}$ , together with plots similar to the one in Fig. 3-right, can be used to benchmark other designs. For instance, the information and regret values obtained for  $\underline{\xi}$  based on covariate-adjusted odds ratio (see, e.g., Hu and Rosenberger (2006, Chap. 9)), with  $[d\xi_1/d\mu](x)$  proportional to  $\eta_1(x)[1 - \eta_2(x)]/\{\eta_2(x)[1 - \eta_1(x)]\}$ , is indicated by a star, showing that it can be improved both in terms of information and regret. The same is true for other rules which are not targeting any specific compromise, in particular those obtained as limits of sequential ad-hoc allocation rules. For instance, one may consider the limits of the information and regret values  $\psi(\underline{\xi}_n)$  and  $R(\underline{\xi}_n)$  (obtained by simulation) for the following generalization of the sequential compromise rule of Hu et al. (2015), which allocates the  $(n + 1)$ -st subject to treatment 1 with probability

$$\pi_1(X_{n+1}) = \frac{d_1^a(\underline{\xi}_n, X_{n+1})}{[1 - \eta_1(X_{n+1})]^b} \left( \sum_{k=1}^2 \frac{d_k^a(\underline{\xi}_n, X_{n+1})}{[1 - \eta_k(X_{n+1})]^b} \right)^{-1}, \quad (2.16)$$

with  $a$  and  $b$  some positive constants,  $d_k^a(\underline{\xi}_n, x) = \text{trace}[\mathbf{M}^{-1}(\underline{\xi}_n)\mathfrak{M}_k(x)]$ , and where  $\underline{\xi}_n$  denotes the empirical design  $(\xi_{n,1}, \xi_{n,2})$ , see (2.3). Taking  $b = 1$  and  $a = 1$  or 2 as suggested in (Hu et al., 2015) gives limiting designs close to  $\underline{\xi}_\mu$ , and

a large regret close to  $R(\underline{\xi}_\mu) \simeq 0.1814$ . For  $a = 1$ , the limiting designs approach  $\underline{\xi}^{*(1)}$  as  $b$  increases, the values for  $b = 3, 4, 6$  and  $10$  are indicated by triangles on Fig. 3-right, from right ( $b = 3$ ) to left ( $b = 10$ ).

In general, the limiting designs for such ad'hoc rules do not have the particular form of optimal designs in Theorem 2.1 and are therefore suboptimal, both in terms of regret and information. Sequential rules that converge to an optimal  $\underline{\xi}^{*(\alpha)}$  for any given  $\alpha$  will be presented in Section 4.

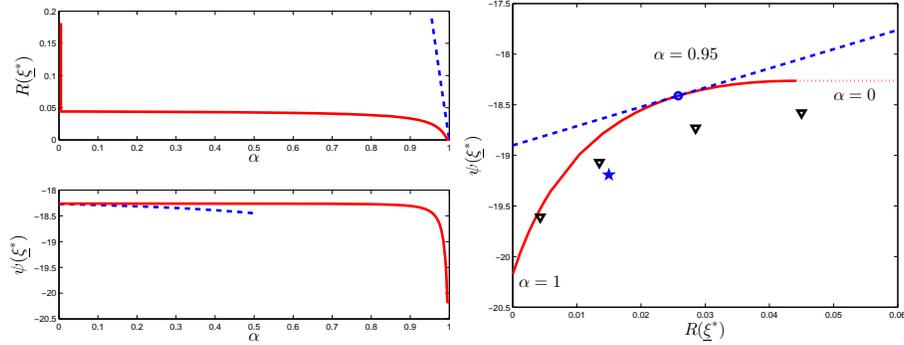


FIG 3. *Example 3* — Left-top:  $R(\underline{\xi}^{*(\alpha)})$  (solid line) and upper bound  $4(1-\alpha)/\alpha$  (dashed line); left-bottom:  $\psi(\underline{\xi}^{*(\alpha)})$  (solid line) and lower bound (2.15) (dashed line); right:  $\psi(\underline{\xi}^{*(\alpha)})$  as a function of  $R(\underline{\xi}^{*(\alpha)})$ ,  $\alpha \in (0, 1]$  with a circle at  $\alpha = 0.95$ ;  $R(\underline{\xi}^{*(\alpha)})$  is not uniquely defined at  $\alpha = 0$ ; the star corresponds to allocation based on covariate-adjusted odds ratio, the triangles to the limiting values for the ad'hoc compromise rule (2.16) with  $a = 1$  and  $b = 3, 4, 6, 10$  from right to left.

## 2.6. Guaranteed minimal allocation proportions

In order to avoid strongly imbalanced allocation, one may wish to set some lower bounds on allocation proportions. In this section we impose that  $\xi_k(\mathcal{X}) \geq \beta/K$  for all  $k$ , with  $\beta \in [0, 1]$  (remember that  $\mu(\mathcal{X}) = 1$ ). An optimal unconstrained design as considered in Section 2.2 may then remain optimal within this framework when  $\beta$  is small enough, but the constraints on allocation proportions modify the characterization of optimal designs for large  $\beta$ .

For any  $\underline{\xi} = (\xi_1, \dots, \xi_K) \in \Xi(\mu)$ , denote

$$\underline{\rho}(\underline{\xi}) = K \min_{k=1, \dots, K} \xi_k(\mathcal{X}). \quad (2.17)$$

The set  $\Xi_\beta(\mu)$  of designs  $\xi \in \Xi(\mu)$  such that  $\underline{\rho}(\xi) \geq \beta$  is convex, and  $\underline{\xi}_P^{*(\alpha, \beta)} \in \Xi_\beta(\mu)$  maximizes  $H^{(\alpha)}(\cdot)$  if and only if  $F_{H^{(\alpha)}}(\underline{\xi}_P^{*(\alpha, \beta)}; \nu) \leq 0$  for all  $\nu \in \Xi_\beta(\mu)$ . In the case  $K = 2$ , this yields the following modification of Corollary 2.1.

**Theorem 2.2** *Suppose that  $K = 2$  and  $H^{(\alpha)}(\cdot)$  is differentiable at  $\underline{\xi}_P^{*(\alpha,\beta)} = \underline{\xi}^* = (\xi_1^*, \xi_2^*) \in \Xi_\beta(\mu)$  for some  $\alpha, \beta \in [0, 1]$ . The following statements are equivalent:*

- (i)  $\underline{\xi}^*$  is optimal, i.e., it maximizes  $H^{(\alpha)}(\underline{\xi})$  with respect to  $\underline{\xi} \in \Xi_\beta(\mu)$ ;
- (ii) there exists a constant  $c = c(\beta)$  such that  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) \geq c$   $\xi_1^*$ -a.e. and  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) \leq c$   $\xi_2^*$ -a.e.;
- (iii) there exist two subsets  $\mathcal{X}_1 = \mathcal{X}_1^{(\alpha,\beta)}$  and  $\mathcal{X}_2 = \mathcal{X}_2^{(\alpha,\beta)}$  of  $\mathcal{X}$  and a constant  $c = c(\beta)$  such that
  - (a)  $\xi_1^* = \mu$  on  $\mathcal{X}_1$  and  $\xi_2^* = \mu$  on  $\mathcal{X}_2$ ,
  - (b)  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) = c$  on  $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$ ,
  - (c)  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) > c$  for all  $x \in \mathcal{X}_1$  and  $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) < c$  for all  $x \in \mathcal{X}_2$ .

The proof is given in Appendix. Note that when  $\mu$  has a density with respect to the Lebesgue measure and  $\beta > 0$ , it is reasonable to assume that  $\mathbf{M}(\underline{\xi})$  has full rank for all  $\underline{\xi} \in \Xi_\beta(\mu)$ , which guarantees the differentiability of  $H^{(\alpha)}(\cdot)$  at  $\underline{\xi}_P^{*(\alpha,\beta)}$ .

Developments similar to those in Section 2.4 show that an optimal design  $\underline{\xi}_P^* = \underline{\xi}_P^{*(\alpha,\beta)} = (\xi_1^*, \xi_2^*)$  satisfies

$$\begin{aligned} \eta_2(x) - \eta_1(x) &\leq \frac{1-\alpha}{\alpha} \text{trace}\{\nabla\psi(\underline{\xi}_P^*)[\mathfrak{M}_1(x) - \mathfrak{M}_2(x)]\} - \frac{c(\beta)}{\alpha}, \quad x \in \mathcal{X} \setminus \mathcal{X}_2 \\ \eta_1(x) - \eta_2(x) &\leq \frac{1-\alpha}{\alpha} \text{trace}\{\nabla\psi(\underline{\xi}_P^*)[\mathfrak{M}_2(x) - \mathfrak{M}_1(x)]\} + \frac{c(\beta)}{\alpha}, \quad x \in \mathcal{X} \setminus \mathcal{X}_1 \end{aligned}$$

Therefore,  $R(\underline{\xi}_P^*) \leq (1-\alpha)/\alpha \text{trace}\{\nabla\psi(\underline{\xi}_P^*)\mathbf{M}(\underline{\xi}_P^*)\} + [\xi_2^*(\mathcal{X}) - \xi_1^*(\mathcal{X})]c(\beta)/\alpha$ , where the second term is non-positive. The comparison between  $\underline{\xi}_P^{*(\alpha,\beta)}$  and  $\underline{\xi}_P^{*(1,\beta)}$  yields an inequality similar to (2.14). Also, the lower bound on information obtained in Section 2.4 remains valid when we consider designs  $\underline{\xi} \in \Xi_\beta(\mu)$ . In particular, the comparison with  $\psi(\underline{\xi}_\mu)$  obtained for balanced random allocation gives  $\psi(\underline{\xi}_P^{*(\alpha,\beta)}) \geq \psi(\underline{\xi}_\mu) - \alpha R(\underline{\xi}_\mu)/(1-\alpha)$ .

### 3. Allocation with randomization

Selection bias occurs if the experimenter is able to correctly guess next allocation in a sequential trial, see, e.g., Rosenberger and Lachin (2002, Chap. 6). The bias factor  $B_n$  for  $n$  allocations is

$$B_n = \frac{\text{nb. of correctly guessed allocations} - \text{nb. of incorret guesses}}{n}.$$

Suppose that allocations follow the optimal design of Theorem 2.1, with allocation to treatment  $j$  when  $x \in \mathcal{X}_j$ . Then, in all situations where condition

(2.12) is satisfied, the bias factor is 1. This is also the case, asymptotically, in a response-adaptive framework where  $\mathcal{X}_j = \mathcal{X}_j(n)$  evolves with the number  $n$  of allocations and depends on the current estimated value  $\hat{\theta}^n$ : the experimenter may still know the rule and predict the next allocation. In this section, we introduce randomization in allocations in the following way: for each subject, with probability  $\beta$  we use random balanced allocation and with probability  $1 - \beta$  we use an optimized, predictable, rule. Then, only the fraction  $1 - \beta$  of allocations can be guessed correctly with certainty, and  $B_n \xrightarrow{\text{a.s.}} 1 - \beta$  as  $n \rightarrow \infty$ . Developments similar to those of Section 2 can be made within this randomized framework and are presented below.

### 3.1. Optimal design and Equivalence Theorem

For any  $\underline{\xi} \in \Xi(\mu)$ , we define the *uniform randomization factor* of  $\underline{\xi}$  as

$$r(\underline{\xi}) = K \min_{k=1, \dots, K} \inf_{x \in \mathcal{X}} \frac{d\xi_k}{d\mu}(x), \quad (3.1)$$

with  $d\xi_k/d\mu$  the Radon-Nikodým derivative of  $\xi_k$  with respect to  $\mu$ . Then, if the  $\xi_{n,j}$  defined by (2.3) tend to  $\xi_j$  as  $n$  tends to infinity (weak convergence),  $B_n$  satisfies  $\limsup_{n \rightarrow \infty} B_n \leq 1 - r(\underline{\xi})$ . Note that  $r(\underline{\xi}) \leq \underline{\rho}(\underline{\xi})$  defined by (2.17), with equality if and only if  $\xi_k = \mu/K$  for some  $k$ .

Consider the maximization of  $H^{(\alpha)}(\underline{\xi})$  with respect to  $\underline{\xi} \in \Xi(\mu)$  under the constraint  $r(\underline{\xi}) \geq \beta$ , for some given  $\beta \in [0, 1]$ . For any  $\underline{\xi} \in \Xi(\mu)$ , the fraction of  $\mu$  that corresponds to allocation of treatment  $k$  can be decomposed as  $\xi_k = (\beta/K)\mu + \tilde{\xi}_k$ , where  $\tilde{\underline{\xi}} = (\tilde{\xi}_1, \dots, \tilde{\xi}_K)$  belongs to  $\Xi[(1 - \beta)\mu]$ , see (2.7). This optimal design problem thus consists in maximizing  $H^{(\alpha, \beta)}(\tilde{\underline{\xi}}) = H^{(\alpha)}(\tilde{\underline{\xi}} + \beta\underline{\xi}_{\underline{\mu}})$  with respect to  $\tilde{\underline{\xi}} \in \Xi[(1 - \beta)\mu]$ , with  $\underline{\xi}_{\underline{\mu}} = (\mu/K, \dots, \mu/K) \in \Xi(\mu)$ .

As in Section 2.2, this is a concave optimization problem over a convex set. We shall denote by  $\underline{\xi}_R^{*(\alpha, \beta)}$  a design in  $\Xi(\mu)$  that maximizes  $H^{(\alpha)}(\underline{\xi})$  under the constraint  $r(\underline{\xi}) \geq \beta$ , for given  $\alpha, \beta \in [0, 1]$ . Such optimal designs  $\underline{\xi}_R^{*(\alpha, \beta)} = \tilde{\underline{\xi}}^* + \beta\underline{\xi}_{\underline{\mu}}$  are still characterized by Theorem 2.1, with the following slight modifications: the statement in (ii) is now valid  $\tilde{\xi}_i^*$ -a.e.;  $\sum_{i \in \mathcal{J}_t} \tilde{\xi}_i^* = (1 - \beta)\mu$  in (iii-a). When  $K = 2$ , Corollary 2.1 is modified as follows: the statements in (ii) are valid  $\tilde{\xi}_1^*$  and  $\tilde{\xi}_2^*$ -a.e.; in (iii),  $\xi_1^* = (1 - \beta/2)\mu$  on  $\mathcal{X}_1$  and  $\xi_2^* = (1 - \beta/2)\mu$  on  $\mathcal{X}_2$ . Note that  $H^{(\alpha, \beta)}(\cdot)$  is differentiable at any  $\tilde{\underline{\xi}} \in \Xi[(1 - \beta)\mu]$  when  $\beta > 0$ , since we have assumed that  $\mathbf{M}(\underline{\xi}_{\underline{\mu}}) \in \mathbb{M}^+$ . The case of unbalanced randomization could be treated in the same way.

### 3.2. Bounds on optimal regret and information

When  $\beta > 0$ , due to linearity in  $\underline{\xi}$  of  $R(\underline{\xi})$ , the regret for an optimal design  $\underline{\xi}^* = \underline{\xi}_R^{*(\alpha, \beta)} = \tilde{\underline{\xi}}^* + \beta\underline{\xi}_{\underline{\mu}}$  can be decomposed into two parts,  $R(\underline{\xi}^*) = R(\tilde{\underline{\xi}}^*) + \beta R(\underline{\xi}_{\underline{\mu}})$ ,

where  $R(\underline{\xi}^*)$  satisfies an inequality similar to (2.13). Therefore

$$R(\underline{\xi}^*) \leq \frac{1-\alpha}{\alpha} \text{trace}[\nabla\psi(\underline{\xi}^*)\mathbf{M}(\underline{\xi}^*)] + \beta R(\underline{\xi}_\mu).$$

When  $\Psi(\mathbf{M}) = \log \det(\mathbf{M})$ , we obtain

$$R(\underline{\xi}^*) \leq p \frac{1-\alpha}{\alpha} + \beta R(\underline{\xi}_\mu) - \beta \frac{1-\alpha}{\alpha} \text{trace}[\mathbf{M}^{-1}(\underline{\xi}^*)\mathbf{M}(\underline{\xi}_\mu)].$$

Since  $\text{trace}[\mathbf{M}^{-1}(\underline{\xi}^*)\mathbf{M}(\underline{\xi}_\mu)] > \text{trace}\{[K\mathbf{M}(\underline{\xi}_\mu)]^{-1}\mathbf{M}(\underline{\xi}_\mu)\} = p/K$ , we have  $R(\underline{\xi}^*) < p(1-\beta/K)(1-\alpha)/\alpha + \beta R(\underline{\xi}_\mu)$ . We can also obtain an inequality similar to (2.14), based on the comparison between  $\underline{\xi}_R^{*(\alpha,\beta)}$  and  $\underline{\xi}_R^{*(1,\beta)}$ .

When  $\alpha < 1$ , we can compare the information  $\psi(\underline{\xi}_R^{*(\alpha,\beta)})$  with that obtained for another design  $\underline{\xi} \in \Xi(\mu)$  such that  $r(\underline{\xi}) \geq \beta$ . As in Section 2.4, we have  $\psi(\underline{\xi}_R^{*(\alpha,\beta)}) \geq \psi(\underline{\xi}) + \alpha[\phi(\underline{\xi}) - \phi(\underline{\xi}_R^{*(\alpha,\beta)})]/(1-\alpha)$  and therefore

$$\begin{aligned} \psi(\underline{\xi}_R^{*(\alpha,\beta)}) &\geq \psi(\underline{\xi}) + \frac{\alpha}{1-\alpha} [\phi(\underline{\xi}) - \max_{\underline{\nu} \in \Xi(\mu): r(\underline{\nu}) \geq \beta} \phi(\underline{\nu})] \\ &= \psi(\underline{\xi}) + \frac{\alpha}{1-\alpha} [\phi(\underline{\xi}) - \beta\phi(\underline{\xi}_\mu) - (1-\beta) \max_{\underline{\nu} \in \Xi(\mu)} \phi(\underline{\nu})], \end{aligned}$$

where we have used the linearity of  $\phi(\underline{\xi})$  with respect to  $\underline{\xi}$ . Using (2.11), we obtain in particular  $\psi(\underline{\xi}_R^{*(\alpha,\beta)}) \geq \psi(\underline{\xi}_\mu) - (1-\beta)R(\underline{\xi}_\mu)\alpha/(1-\alpha)$ , to be compared with (2.15).

### 3.3. Example 4

We modify the allocation problem in Example 3, and take now  $\eta_1(x) = \eta_1(x, \theta_1) = 0.1 + 0.5 e^{z_1(x)}/(1 + e^{z_1(x)})$ . We also introduce balanced random allocation through the constraint  $r(\underline{\xi}) \geq \beta = 0.2$ . The optimal designs obtained for  $\alpha \in [0, 1]$  are presented in Fig. 4-right. Note that there is a range of values of  $\alpha$  for which the sets  $\mathcal{X}_j^{(\alpha)}$  are now the unions of three intervals, compare with Fig. 2-right. Also, for  $\alpha = 0$  and all  $\beta \in [0, 1]$  the optimal designs  $\underline{\xi}_R^{*(0,\beta)}$  are now uniquely defined.

## 4. Covariate-adaptive sequential allocation targeting an optimal design

Denote by  $\mathbf{T}_1, \mathbf{T}_2, \dots$  the sequence of treatment assignments, where  $\mathbf{T}_n = (T_{n,1}, \dots, T_{n,K})$  with  $T_{n,j} = 1$  when the  $n$ -th subject, with covariates  $X_n$ , is allocated to treatment  $j$ , all  $T_{n,i}$  with  $i \neq j$  being then zero. The  $(n+1)$ -st subject is allocated to treatment  $k$  with probability

$$\pi_k(X_{n+1}) = \text{Prob}(T_{n+1,k} = 1 | X_{n+1}, \mathcal{F}_n), \quad k = 1, \dots, K, \quad (4.1)$$

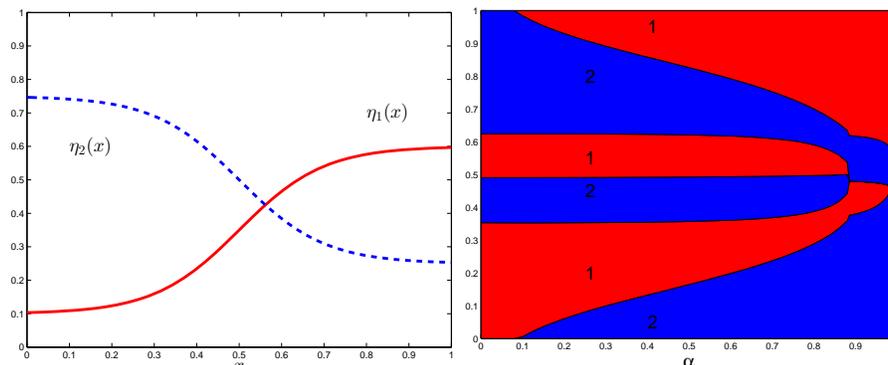


FIG 4. *Example 2* — Left:  $\eta_1(x)$  (solid line) and  $\eta_2(x)$  (dashed line) as functions of  $x$ ; right: subsets  $\mathcal{X}_1^{(\alpha)}$  and  $\mathcal{X}_2^{(\alpha)}$  of Corollary 2.1-(ii) for  $\alpha \in [0, 1]$  and  $\beta = 0.2$ .

with  $\mathcal{F}_n$  the sigma field  $\sigma(\mathbf{T}_1, \dots, \mathbf{T}_n, X_1, \dots, X_n)$ . Also denote  $\mathbf{N}_n = \sum_{i=1}^n \mathbf{T}_i$ , so that its  $k$ -th component  $N_{n,k}$  is the number of subjects allocated to treatment  $k$  in the first  $n$  assignments. The empirical measures  $\xi_{n,k}$  defined in (2.3) are then given by  $\xi_{n,k} = (1/n) \sum_{i=1}^n T_{i,k} \delta_{X_i}$ .

In this section we present different choices for  $\pi_k(\cdot)$  in (4.1) that asymptotically achieve the limiting allocation given by one of the optimal designs  $\underline{\xi}^* = (\xi_1^*, \dots, \xi_K^*)$  considered in Sections 2 and 3. A first easy solution to achieve  $\xi_{n,k} \xrightarrow{\text{a.s.}} \xi_k^*$  (weak convergence) is to sample according to  $\xi_k^*$  in (4.1). This is considered in Section 4.1. A second rule, design adaptive in the sense that  $\pi_k(X_{n+1})$  depends on the  $\xi_{n,k}$ ,  $k = 1, \dots, K$ , is considered in Section 4.2. We always assume that the first  $n_0$  subjects are allocated with some predefined rule (e.g., balanced random allocation), for some  $n_0 > 0$ .

#### 4.1. Sequential allocation based on oracle optimal design

Consider the sequential allocation rule defined by

$$\pi_k^*(X_{n+1}) = \frac{d\xi_k^*}{d\mu}(X_{n+1}), \quad n \geq n_0, \quad (4.2)$$

where  $\underline{\xi}^*$  denotes an optimal design as in Section 2, or an optimal design  $\underline{\xi}_R^{*(\alpha, \beta)}$  satisfying  $r(\underline{\xi}_R^{*(\alpha, \beta)}) \geq \beta$  as in Section 3, see (3.1). In particular, when (2.12) is satisfied, then (4.2) simply corresponds to

$$\pi_k^*(X_{n+1}) = \begin{cases} 1 & \text{if } G_k^{(\alpha)}(\underline{\xi}^*, X_{n+1}) = \max_{j=1, \dots, K} G_k^{(\alpha)}(\underline{\xi}^*, X_{n+1}) \\ 0 & \text{otherwise} \end{cases}$$

when  $\beta = 0$ , and more generally to

$$\pi_k^*(X_{n+1}) = \begin{cases} 1 - (K-1)\beta/K & \text{if } G_k^{(\alpha)}(\underline{\xi}^*, X_{n+1}) = \max_{j=1, \dots, K} G_j^{(\alpha)}(\underline{\xi}^*, X_{n+1}) \\ \beta/K & \text{otherwise} \end{cases} \quad (4.3)$$

for  $\beta \geq 0$ .  $\mathbf{T}_n$  follows a multinomial distribution, with  $\text{Prob}\{T_{n,k} = 1\} = \xi_k^*(\mathcal{X})$  for all  $k$  and  $n$ . Therefore,

$$\mathbf{N}/n \xrightarrow{\text{a.s.}} \rho^* = \rho(\underline{\xi}^*) = (\xi_1^*(\mathcal{X}), \dots, \xi_K^*(\mathcal{X}))^\top \quad (4.4)$$

from the Strong Law of Large Numbers (SLLN), and  $\sqrt{n}(\mathbf{N}/n - \rho^*) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma^*)$  with  $\Sigma^* = \text{diag}(\rho^*) - \rho^* \rho^{*\top}$  from the Central Limit Theorem (CLT). The proportions  $N_n/n$  also satisfy the Law of the Iterated Logarithm (LIL), with  $\limsup_{n \rightarrow \infty} \sqrt{n/(2\sigma_k^2 \text{Log Log } n)} |N_{n,k}/n - \rho_k^*| = 1$  (a.s.) for all  $k$ , where  $\sigma_k^2 = \Sigma_{kk}^*$  and  $\text{Log } t = \max\{1, \log t\}$  for all  $t > 0$ . Moreover,  $R_n = R(\underline{\xi}_n^*)$  and  $\mathbf{M}_n = \mathbf{M}(\underline{\xi}_n^*)$  respectively given by (2.2) and (2.5) satisfy  $R_n \xrightarrow{\text{a.s.}} R(\underline{\xi}^*)$  and  $\mathbf{M}_n \xrightarrow{\text{a.s.}} \mathbf{M}(\underline{\xi}^*)$ , where  $\underline{\xi}_n^* = (\xi_{n,1}^*, \dots, \xi_{n,K}^*)$  with  $\xi_{n,k}^*$  the empirical measure (2.3). We thus have  $\psi(\underline{\xi}_n^*) \xrightarrow{\text{a.s.}} \psi(\underline{\xi}^*)$  and  $H^{(\alpha)}(\underline{\xi}_n^*) \xrightarrow{\text{a.s.}} H^{(\alpha)}(\underline{\xi}^*)$ . The values of  $R(\underline{\xi}_n^*)$  and  $\psi(\underline{\xi}_n^*)$  also obey the CLT; direct calculations show that  $\sqrt{n}[R(\underline{\xi}_n^*) - R(\underline{\xi}^*)] \xrightarrow{d} \mathcal{N}(0, V_R^*)$  and (using the delta method) that  $\sqrt{n}[\psi(\underline{\xi}_n^*) - \psi(\underline{\xi}^*)] \xrightarrow{d} \mathcal{N}(0, V_\psi^*)$ , with

$$V_R^* = \sum_{k=1}^K \mathbb{P}_{\xi_k^*} (\eta_k^* - \eta_k)^2 - R^2(\underline{\xi}^*), \quad (4.5)$$

$$V_\psi^* = \sum_{k=1}^K \mathbb{P}_{\xi_k^*} \text{trace}^2[\nabla \psi(\underline{\xi}^*) \mathfrak{M}_k(\cdot)] - \text{trace}^2[\nabla \psi(\underline{\xi}^*) \mathbf{M}(\underline{\xi}^*)]. \quad (4.6)$$

Although attractive from a theoretical viewpoint, (4.2) has the inconvenient that it relies on the prior construction of an optimal design  $\underline{\xi}^*$ . In particular, extension to response-adaptive allocation may be unpractical: indeed, allocation of the  $(n+1)$ -st subject should then be based on the optimal design  $\underline{\xi}^*(\hat{\theta}^n)$  for the current estimated value  $\hat{\theta}^n$  of  $\theta$ , see Section 5, which means that an oracle providing  $\underline{\xi}^*(\theta)$  for any  $\theta$  should be available. In the next section we consider an allocation rule  $\hat{\pi}_k(X_{n+1})$  that asymptotically samples from  $\underline{\xi}^*$  without requiring neither the explicit construction of  $\underline{\xi}^*$  nor the knowledge of  $\mu$ .

#### 4.2. Doubly-adaptive sequential allocation

The rule is based on the substitution of  $\hat{\underline{\xi}}_n = (\hat{\xi}_{n,1}, \dots, \hat{\xi}_{n,K})$  for  $\underline{\xi}^*$  in (4.3), with  $\hat{\xi}_{n,k}$  the empirical measure (2.3) for the sequential assignments. It is covariate and design-adaptive, i.e., adaptive also with respect to previous allocations, and

uses allocation probabilities given by

$$\hat{\pi}_k(X_{n+1}) = \begin{cases} \frac{1 - [K - \ell_n(X_{n+1})]\beta/K}{\ell_n(X_{n+1})} & \text{if} \\ \beta/K & \text{otherwise,} \end{cases} \quad G_k^{(\alpha)}(\hat{\underline{\xi}}_n, X_{n+1}) = \max_{j=1, \dots, K} G_j^{(\alpha)}(\hat{\underline{\xi}}_n, X_{n+1}) \quad (4.7)$$

where  $\ell_n(x) = \left| \left\{ j \in \{1, \dots, K\} : G_j^{(\alpha)}(\hat{\underline{\xi}}_n, x) = \max_{k=1, \dots, K} G_k^{(\alpha)}(\hat{\underline{\xi}}_n, x) \right\} \right|$ . When  $\ell_n(X_{n+1}) = 1$ , then  $\hat{\pi}_k(X_{n+1}) = 1 - (K-1)\beta/K$  for  $k$  such that  $G_k^{(\alpha)}(\hat{\underline{\xi}}_n, X_{n+1}) = \max_{j=1, \dots, K} G_j^{(\alpha)}(\hat{\underline{\xi}}_n, X_{n+1})$ .

Theorem 4.1 indicates that when  $\beta > 0$ ,  $\hat{\underline{\xi}}_n$  generated by (4.7) has the same asymptotic information and regret values as  $\underline{\xi}_R^{*(\alpha, \beta)}$  of Section 3. When  $\alpha = 0$  and  $\Psi(\mathbf{M}) = \log \det(\mathbf{M})$ , (4.7) corresponds to the sequential construction of a  $D$ -optimal design in  $\Xi(\mu)$ , see Atkinson (1982, 1999, 2002). Notice, however, that the investigation of the convergence properties of such extensions of biased-coin designs with covariate information, based on optimal design theory, has received little attention in the literature, if any. Also note that the allocation rule (4.7) does not enter the general framework considered in (Baldi Antognini and Zagoraiou, 2015). The proof of Theorem 4.1 is given in Appendix. The assumption that  $\mathcal{X}$  is bounded in H1b can be relaxed, at the expense of adding suitable moment conditions on  $\mathfrak{M}_k(X)$  and growth condition on  $\nabla^2 \Psi(\cdot)$  to H2 and H2', similarly to Pronzato (2006, Th. 9).

**Theorem 4.1** *Under H1a or (H1b, H2, H2'), for any  $\beta \in (0, 1]$  and  $\alpha \in [0, 1]$ , the allocation rule (4.7) satisfies*

$$H^{(\alpha)}(\hat{\underline{\xi}}_n) \xrightarrow{\text{a.s.}} H^{(\alpha)}(\underline{\xi}_R^{*(\alpha, \beta)}), \quad n \rightarrow \infty,$$

with  $\underline{\xi}_R^{*(\alpha, \beta)}$  an optimal design maximizing  $H^{(\alpha)}(\underline{\xi})$  with respect to  $\underline{\xi} \in \Xi(\mu)$  under the constraint  $r(\underline{\xi}) \geq \beta$ , see (3.1). Moreover,  $\mathbf{M}(\hat{\underline{\xi}}_n) \xrightarrow{\text{a.s.}} \mathbf{M}(\underline{\xi}_R^{*(\alpha, \beta)})$ ,  $\psi(\hat{\underline{\xi}}_n) \xrightarrow{\text{a.s.}} \psi(\underline{\xi}_R^{*(\alpha, \beta)})$ , and also  $R(\hat{\underline{\xi}}_n) \xrightarrow{\text{a.s.}} R(\underline{\xi}_R^{*(\alpha, \beta)})$  if  $\alpha > 0$ .

When  $\alpha < 1$ , the assumption that  $\beta > 0$  permits to bound the second-order derivative of  $\psi(\cdot)$  from below and is crucial in the proof of the theorem. For  $\beta = 0$ , we only have a dichotomous property, similar to that in (Wu and Wynn, 1978): either  $H^{(\alpha)}(\hat{\underline{\xi}}_n) \rightarrow H^{(\alpha)}(\underline{\xi}_R^{*(\alpha, \beta)})$ , or  $\liminf_{n \rightarrow \infty} \Psi[\mathbf{M}(\hat{\underline{\xi}}_n)] = -\infty$  when  $\Psi(\cdot)$  is one of the criteria (2.6) ( $\liminf_{n \rightarrow \infty} \Psi[\mathbf{M}(\hat{\underline{\xi}}_n)] = 0$  for their positively homogeneous versions). However, the  $X_i$  being i.i.d. in  $\mathcal{X}$  with  $\mathbf{M}(\underline{\xi}_\mu) \in \mathbb{M}^+$  (where  $\xi_{\mu_k} = \mu/K$  for all  $k$ ), one can force the second event to have zero probability. For  $\alpha \in [0, 1)$  and  $\beta = 0$ , we modify the rule (4.7) through the introduction of a lower bound  $\underline{\psi}^{(\alpha)}$  on  $\psi(\underline{\xi}_R^{*(\alpha)})$ , obtained for instance from (2.15). For each  $n \geq n_0$ , we allocate the  $(n+1)$ -st subject to treatment  $k$  with probability  $1/K$  if  $\psi(\hat{\underline{\xi}}_n) < \min\{\underline{\psi}^{(\alpha)}, \psi(\underline{\xi}_\mu)\}$  and with probability  $\hat{\pi}_k^{\beta=0}(X_{n+1})$  otherwise, where  $\hat{\pi}_k^{\beta=0}(X_{n+1})$  substitutes  $\beta = 0$  in (4.7). Then, for  $\alpha < 1$  and

$\beta = 0$ , the second-order derivative of  $\psi(\cdot)$  at  $\hat{\xi}_n$  is bounded from below (a.s.), and the empirical design  $\hat{\xi}_n$  obtained with this modified allocation rule satisfies the same asymptotic properties as in Theorem 4.1.

From Theorem 4.1, when  $\alpha, \beta > 0$  (or  $\beta = 0$  with  $0 < \alpha < 1$  for the modified rule just above), the information and regret values obtained with (4.7) converge (a.s.) to those obtained with the rule (4.2) based on an oracle optimal design. Under the conditions mentioned in Section 2.3.3, this is also the case for the allocation proportions  $\mathbf{N}/n$ . On the other hand, numerical simulations indicate that their asymptotic variance  $\widehat{\Sigma}$  is smaller than  $\Sigma^*$  obtained with  $\pi_k^*(X_{n+1})$ , a phenomenon that resembles the improved treatment balance obtained by the method of Pocock and Simon (1975) generalizing Efron (1971). The doubly-adaptive designs of Zhang and Hu (2009), which extend the approach of Hu and Zhang (2004) to the presence of covariates, are able to yield (in the limit) a reduction of  $\Sigma^*$  to

$$\text{Var}(\underline{\pi}^*) = \mathbf{P}_\mu(\underline{\pi}^*(\cdot) - \rho^*)(\underline{\pi}^*(\cdot) - \rho^*)^\top,$$

where  $\underline{\pi}^*(x) = (\pi_1^*(x), \dots, \pi_K^*(x))^\top$  with  $\pi_k^*(x)$  and  $\rho^*$  given by (4.2) and (4.4). When  $\beta = 0$  and (2.12) is satisfied (so that  $\pi_k^*(x) \in \{0, 1\}$  for all  $x$ ), then  $\text{Var}(\underline{\pi}^*) = \Sigma^*$  and numerical simulations show that the rule  $\hat{\pi}_k(X_{n+1})$  given by (4.7) achieves a smaller asymptotic variance than  $\text{Var}(\underline{\pi}^*)$  for the proportions  $\mathbf{N}/n$ , indicating that  $\text{Var}(\underline{\pi}^*)$  is not the best asymptotic variability within the class of covariate-adjusted designs considered in (Zhang and Hu, 2009, Eq. (2.9)). This will be illustrated in Section 4.3.2 which continues Example 3. Note that the approach used in the same paper for the derivation of  $\widehat{\Sigma}$ , based on a functional CLT, seems difficult to extend to our situation where the design adaptation concerns the whole matrix  $\mathbf{M}(\hat{\xi}_n)$  and not only the proportions  $\mathbf{N}/n$ .

When  $\xi^*$  is not unique, one may wonder what is the limiting design for (4.7). Numerical simulations indicate convergence to a unique design, whatever the initialization of the sequential procedure (with  $n_0$  arbitrarily large). Section 4.3.1 gives an illustration through a continuation of Example 1 of Section 2.5.1. Further developments are required to investigate if the stability properties of (4.7) around an optimal design  $\xi^*$  permit to characterize which particular optimal designs can be reached in the limit.

### 4.3. Examples

#### 4.3.1. Example 1 (continued)

Figure 5 presents histograms of  $\hat{\xi}_{n,1}$  and  $\hat{\xi}_{n,2}$  obtained with (4.7) in the situation where there are infinitely many optimal designs  $\xi^*$  ( $\alpha = 0.75 < \underline{\alpha} = 0.96$ , see Section 2.5.1). Whatever the initialization of (4.7), we always observe convergence of  $\hat{\xi}_n$  to the same limiting optimal design.

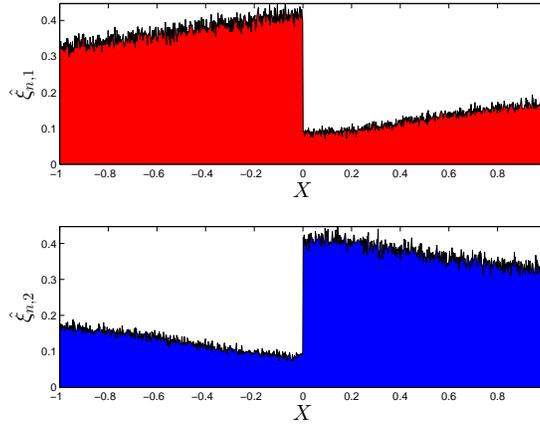


FIG 5. Example 1: histograms of  $\hat{\xi}_{n,1}$  and  $\hat{\xi}_{n,2}$ ,  $n = 10^6$  ( $\beta = 0$ ,  $\alpha = 0.75 < \underline{\alpha} = 0.96$ ).

#### 4.3.2. Example 3 (continued)

Consider again the situation of Example 3, see Section 2.5.3, with  $\alpha = 0.7$  and  $\beta = 0$ . We have performed 1,000 simulations of allocation rules (4.3) and (4.7) with  $n = 5,000$  subjects (we use  $n_0 = 4$ , i.e., two initial assignments of each treatment). Empirical distributions are smoothed with a normal kernel density estimator, using Silverman's rule for bandwidth selection.

Figure 6-left shows the empirical distributions of  $\sqrt{n}(N_{n,k}/n - \rho_k^*)$ , for  $k = 1$  (top) and  $k = 2$  (bottom). The dashed-line curve is for (4.3) and shows good agreement with the asymptotic distribution  $\mathcal{N}(0, \sigma_k^2)$  (solid-line); the dotted-line curve is for (4.7) which exhibits smaller variability around the optimal proportions  $\rho^*$ .

Figure 6-right presents the empirical distributions of  $\sqrt{n}[\psi(\underline{\xi}_n) - \psi(\underline{\xi}^*)]$  (top) and  $\sqrt{n}[R(\underline{\xi}_n) - R(\underline{\xi}^*)]$  (bottom). There is good agreement with the asymptotic distributions  $\mathcal{N}(0, V_\psi^*)$  and  $\mathcal{N}(0, V_R^*)$  (solid lines) for (4.3) (dashed lines), where  $V_R^*$  and  $V_\psi^*$  are given by (4.5) and (4.6). The simulations indicate that the bias on  $R_n$  for the rule (4.7) decreases more slowly than  $1/n$ ; see the dotted line on Fig. 6-bottom-right, together with Fig. 7-left which presents the log of the empirical bias  $\overline{R}_n - R(\underline{\xi}^*)$  (obtained from 1,000 simulations) as a function of  $n$ : the bias is  $O(1/n)$  for (4.3), but between  $O(1/\sqrt{n})$  and  $O(1/n)$  for (4.7). Simulations with other values of  $\alpha$  show that this bias decreases faster as  $\alpha$  tends to 1, see Fig. 7-right which corresponds to  $\alpha = 0.98$ .

#### 4.3.3. Example 4 (continued)

Consider again the situation of Example 4, see Section 3.3, with  $\alpha = 0.7$  and  $\beta = 0.2$ . Figure 8-left presents histograms of  $\hat{\xi}_{n,1}$  and  $\hat{\xi}_{n,2}$  obtained with allocation rule (4.7) (with  $n = 2 \cdot 10^4$  and  $n_0 = 4$  — two initial assignments of each

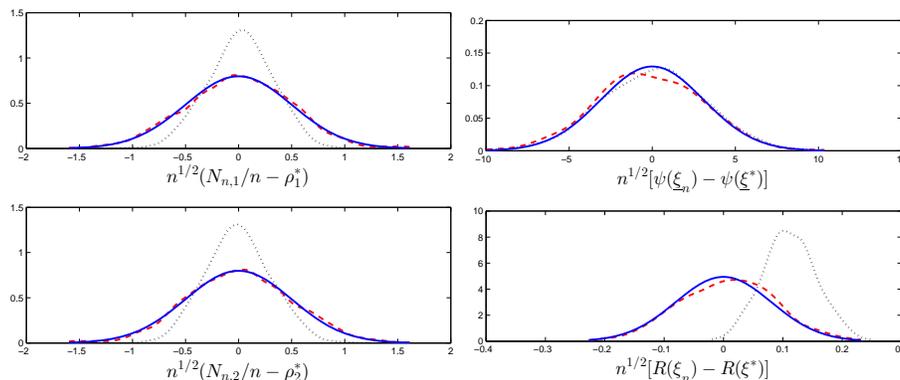


FIG 6. *Example 3* ( $\alpha = 0.7$ ,  $\beta = 0$ ,  $n = 5,000$ , 1,000 repetitions) — *Left*: empirical distributions of  $\sqrt{n}(N_{n,k}/n - \rho_k^*)$ ,  $k = 1, 2$ , for (4.3) (dashed-lines) and (4.7) (dotted lines), the solid-lines correspond to  $\mathcal{N}(0, \sigma_k^2)$ ; *right*: empirical distributions of  $\sqrt{n}[\psi(\underline{\xi}_n) - \psi(\underline{\xi}^*)]$  and  $\sqrt{n}[R(\underline{\xi}_n) - R(\underline{\xi}^*)]$  for (4.3) (dashed-lines) and (4.7) (dotted lines), normal distributions  $\mathcal{N}(0, V_\psi^*)$  and  $\mathcal{N}(0, V_R^*)$  (solid lines).

treatment). Note the good agreement with the optimal design  $\underline{\xi}_R^{*(\alpha, \beta)}$  presented in Fig.4-right, where  $\xi_1^* = (1 - \beta/2)\mu$  on  $\mathcal{X}_1 \simeq (0.237, 0.368) \cup (0.495, 0.610) \cup (0.7525, 1]$ .

In fact, for most assignments the rule (4.7) agrees with (4.3) which samples from  $\underline{\xi}_R^{*(\alpha, \beta)}$ : different allocations occur essentially for values of  $x$  near the endpoints of the intervals that define  $\mathcal{X}_1$ , see Fig. 8-right for a histogram of the values of  $X$  where (4.3) would give a treatment different from that given by (4.7). Figure 9-left shows the number  $N_n^D$  of disagreements with (4.3) when (4.7) is used in a sequence of length  $n$ .  $G_j^{(\alpha)}(\hat{\xi}_n, x)$  converges to  $G_j(\underline{\xi}_R^{*(\alpha, \beta)}, x)$  in  $1/\sqrt{n}$ ,  $j = 1, 2$ , and  $N_n^D$  increases as  $\sqrt{n}$ , see the curve in dashed line. Figure 9-right presents the evolution of  $R(\hat{\xi}_n)$  (top) and  $H(\hat{\xi}_n)$  (bottom) as functions of  $n$  for the rule (4.7). Convergence to the optimal values (indicated by dashed lines) is reasonably fast; the figure is quasi identical when (4.3) is used with the same sequence of covariates.

#### 4.3.4. Example 5

We add a third model to Example 4,  $\eta_3(x, \theta_3) = \theta_3 = 0.1$  (so that  $\eta_3(x, \theta_3) < \min\{\eta_1(x, \theta_1), \eta_2(x, \theta_2)\}$  for all  $x$ , with the third treatment representing for instance placebo). Histograms of  $\hat{\xi}_{j,n}$  obtained with (4.7) ( $n = 2 \cdot 10^4$  and  $n_0 = 6$ ) are presented in Fig. 10-left and right, respectively for  $\alpha = 0.7$ ,  $\beta = 0.2$  and  $\alpha = 0.9$ ,  $\beta = 0.2$ . For large enough  $\alpha$  or  $\beta$ , the optimal design is such that  $\xi_3^* = (\beta/3)\mu$ , i.e., the random component of the design gives enough precision for the estimation of  $\theta_3$  (the placebo effect), taking the poor efficacy of this treatment into account.

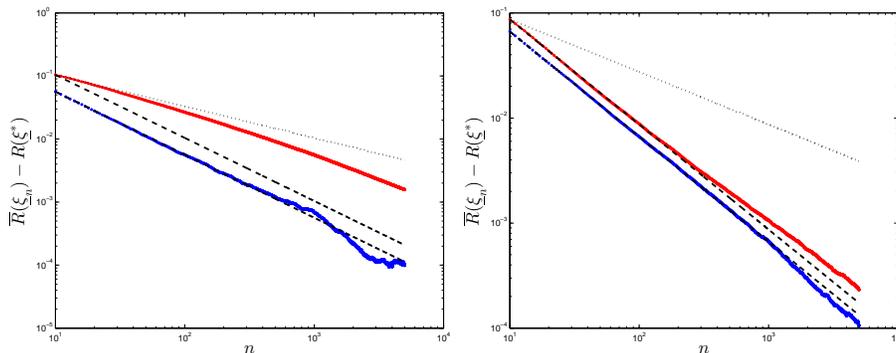


FIG 7. Example 3: empirical bias  $\overline{R}_n - R(\xi^*)$  (1,000 repetitions) as a function of  $n$  (log scale) in solid line for (4.3) (bottom) and (4.7) (top), the dashed lines correspond to a decrease rate  $1/n$ , the dotted line (top) to  $1/\sqrt{n}$ ; left:  $\alpha = 0.7$ , right:  $\alpha = 0.98$ .

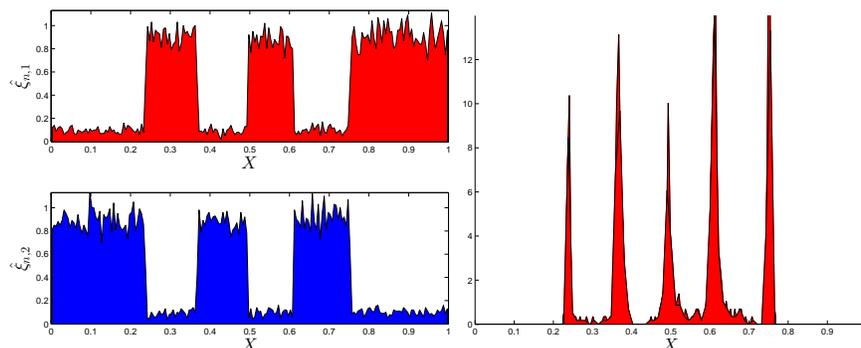


FIG 8. Example 4 ( $\alpha = 0.7$ ,  $\beta = 0.2$ ) — Left: histograms of  $\hat{\xi}_{n,1}$  and  $\hat{\xi}_{n,2}$ ,  $n = 2 \cdot 10^4$ ; right: histogram of covariates for which (4.3) would disagree from (4.7).

## 5. Further extensions and developments

The paper proposes sequential allocation rules that target optimal strategies, in the sense that the asymptotic regret and information values are non dominated, contrasting with ad’hoc rules whose asymptotic regret and information can usually both be improved. These results can be extended in various directions.

**Extension to response-adaptive rules.** As usual in nonlinear situations, optimal designs  $\xi_{\theta}^{*(\alpha)}$  depend on the unknown value  $\theta$  of the model parameters. Here we only considered locally optimum design, where  $\theta$  is set to a given nominal value  $\theta^0$ . In a response-adaptive implementation, when assigning the  $(n + 1)$ -st subject,  $\theta^0$  can be replaced by  $\hat{\theta}^n$ , the current ML estimator of  $\theta$  based on the  $n$  responses observed previously. The asymptotic properties given in Section 4.1 and Theorem 4.1 must be reconsidered when such CARA de-

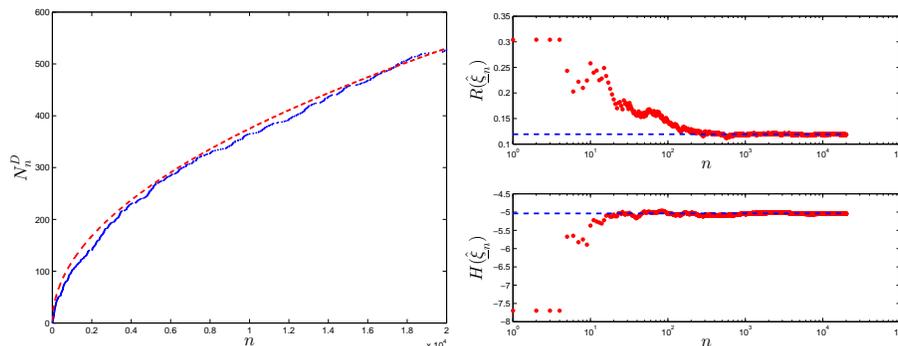


FIG 9. Example 4 ( $\alpha = 0.7$ ,  $\beta = 0.2$ ) — Left: number of disagreements with (4.3) when (4.7) is used in a sequence of length  $n$  (the curve in dashed line corresponds to  $3.75\sqrt{n}$ ); right:  $R(\hat{\xi}_n)$  (top) and  $H(\hat{\xi}_n)$  (bottom) as functions of  $n$  for allocation rule (4.7), optimal values are indicated in dashed lines.

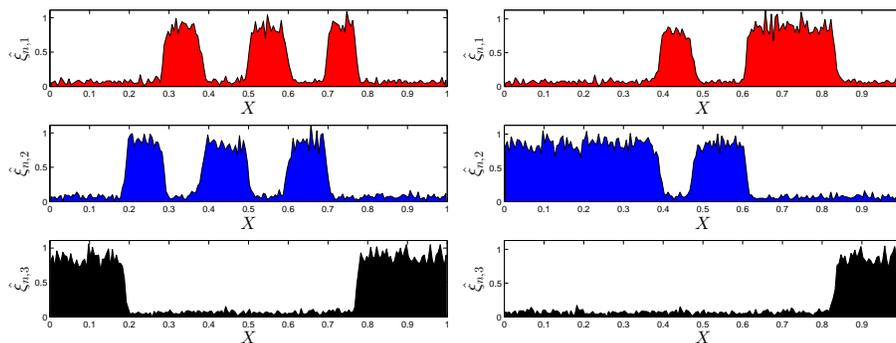


FIG 10. Example 5 — Histograms of  $\hat{\xi}_{n,k}$ ,  $k = 1, 2, 3$  ( $n = 2 \cdot 10^4$ ) for  $\alpha = 0.7$ ,  $\beta = 0.2$  (left) and  $\alpha = 0.9$ ,  $\beta = 0.2$  (right).

signs are used. In particular, the asymptotic variances of the proportions  $\mathbf{N}/n$ , information and regret are modified (increased) compared to Section 4 due to adaptation of allocations to  $\hat{\theta}^n$ . Only a few indications are given below, detailed developments on the asymptotic properties of these CARA designs for generalized linear models, which cover a broad class of applications, will be presented in a forthcoming paper. How to define and obtain an optimal allocation scheme is still considered as an open problem, see Zhang et al. (2007). To our knowledge, this is the first attempt to incorporate covariate information in a response-adaptive design which converges to an optimal target.

First, one may consider a response-adaptive version of (4.2) based on the construction of  $\xi_{\hat{\theta}^n}^{*(\alpha)}$  for each  $n$ , with  $\hat{\theta}^n$  in a compact subset  $\Theta$  of  $\mathbb{R}^p$  containing admissible  $\theta$ . When  $\alpha < 1$ , developments similar to those in (Zhang et al., 2007) can be used to prove the strong consistency and asymptotic nor-

mality of  $\hat{\theta}^n$ , with  $\sqrt{n}(\hat{\theta}^n - \theta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{M}^{-1}(\underline{\xi}_{\theta}^{*(\alpha)}; \theta))$  under rather standard regularity assumptions. Condition A in the same paper is not satisfied since  $\mu\{x : \pi_k^*(x; \theta) = 0\} > 0$  when  $\pi_k^*(x; \theta)$  is given by (4.2), but it maybe counter-balanced by the assumption that all optimal information matrices  $\mathbf{M}(\underline{\xi}_{\theta}^{*(\alpha)}; \theta)$ ,  $\theta \in \Theta$ , are nonsingular. One may also ensure the asymptotic normality of allocation proportions  $\mathbf{N}/n$ , information and regret, like in Section 4.1 but with larger variances. One may try to reduce their values, following the approach in (Zhang and Hu, 2009), provided that the allocation probabilities are suitably smoothed to remove abrupt fronts.

Although theoretically feasible, the computation of  $\underline{\xi}_{\hat{\theta}^n}^{*(\alpha)}$  for each  $n$  is rather inconvenient. One may circumvent that difficulty by constructing approximations  $\widehat{G}_k^{(\alpha)}(\underline{\xi}_{\theta}^{*(\alpha)}, x; \theta)$  of the functions  $(x, \theta) \in (\mathcal{X} \times \Theta) \rightarrow G_k^{(\alpha)}(\underline{\xi}_{\theta}^{*(\alpha)}, x; \theta)$ ,  $k = 1, \dots, K$ : using Theorem 2.1 and the values of  $\widehat{G}_k^{(\alpha)}(\underline{\xi}_{\hat{\theta}^n}^{*(\alpha)}, X_{n+1}; \hat{\theta}^n)$ , one may then compute the  $K$  allocation probabilities  $\widehat{\pi}_k^*(X_{n+1}; \hat{\theta}^n)$  for every  $\hat{\theta}^n$ .

Alternatively, one may consider a response-adaptive version of (4.7), with  $G_k^{(\alpha)}(\underline{\xi}_n, x; \hat{\theta}^n)$  substituted for  $G_k^{(\alpha)}(\underline{\xi}_n, x; \theta^0)$ . The strong consistency of  $\hat{\theta}^n$  and its asymptotic normality  $\sqrt{n}(\hat{\theta}^n - \theta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{M}^{-1}(\underline{\xi}_{\theta}^{*(\alpha)}; \theta))$  can be preserved under suitable regularity assumptions. This is essential since it provides a justification for the use of  $\Psi[\mathbf{M}(\underline{\xi}_{\theta}^{*(\alpha)}; \theta)]$  as a measure of the information content of the experiment.

**Other cumulative regrets.** As an alternative to (2.2) which relies on cumulative treatment responses, one may relate the regret to the number of subjects not receiving the best treatment (i.e., receiving the worst treatment when  $K = 2$ ) and consider  $R_n = (1/n) \sum_{i=1}^n \mathbb{I}\{\eta_{k_i}(X_i) \neq \eta_*(X_i)\} = 1 - \sum_{k=1}^k \xi_{n,k}\{x : \eta_k(x) = \eta_*(x)\}$ . The responses  $\eta_k$  must then be replaced by the indicator functions  $\eta'_k(x) = \mathbb{I}\{\eta_k(x) = \eta_*(x)\}$  in  $\phi(\xi)$ , see (2.9), and  $G_k^{(\alpha)}(\xi, x)$ , see (2.10). Also, one may enforce individual ethics by increasing the penalty for not using the best treatment, and consider  $R_n = \sum_{l=1}^n \mathbb{P}_{\xi_{n,k}}[\eta_k^q(\cdot) - \eta_*^q(\cdot)]$  with  $q < 0$ . Most developments in the paper remain valid since  $R(\xi)$  is still linear in  $\underline{\xi}$ .

**Adaptive choice of  $\alpha$ .** It seems difficult to select a suitable value for  $\alpha$  in absence of information on the performance of the corresponding optimal design  $\underline{\xi}_{\theta}^{*(\alpha)}$ , even if the bounds in Sections 2.4 and 3.2 may help. An alternative solution is to specify a target  $\tau$  on the regret, and maximize information under the constraint that  $R(\xi) \leq \tau$ , with associated Lagrangian  $\mathcal{L}(\xi, C) = \psi(\xi; \theta) - C[R(\xi; \theta) - \tau]$ , see Section 2.1. In a CARA scheme, one may then let the Lagrange coefficient  $C$  vary with  $n$  as  $C_{n+1} = \max\{0, C_n + \gamma[R(\hat{\xi}_n; \hat{\theta}^n) - \tau]\}$ , with  $\gamma$  some positive constant. This is equivalent to letting  $\alpha$  depend on  $n$ , with  $C_n = \alpha_n/(1 - \alpha_n)$  and  $\alpha_{n+1} = C_{n+1}/(1 + C_{n+1})$ .

Finally, one may consider adaptive strategies that give an increasing importance to allocation to the best treatment, and let  $\alpha = \alpha_n$  tend to 1 as  $n \rightarrow \infty$  in a CARA rule. This is equivalent to letting  $C_n = \alpha_n/(1 - \alpha_n)$  tend to infinity, which raises several open questions: which increase rate for  $C_n$  ensures the

strong consistency of  $\hat{\theta}^n$ ? Is it possible to reach the best achievable decrease rate for the expected regret in this context, that is  $\mathbb{E}\{R_n\} = O(\log(n)/n)$ , see for instance Lai and Robbins (1985); Cappé et al. (2013); Goldenshluger and Zeevi (2013)?

## Appendix: proofs

*Proof of Theorem 2.1.* The proof closely follows the one in (Sahm and Schwabe, 2001).

Proof that (i)  $\implies$  (ii). Suppose that (ii) is not satisfied. It implies that there exists  $\mathcal{A} \subset \mathcal{X}$  with  $\xi_i^*(\mathcal{A}) > 0$  and  $G_i^{(\alpha)}(\underline{\xi}^*, x) < G_j^{(\alpha)}(\underline{\xi}^*, x)$  for some  $j \neq i$  and all  $x \in \mathcal{A}$ . Consider the measure  $\underline{\nu} \in \Xi(\mu)$  defined by  $\underline{\nu} = \underline{\xi}^*$  on  $\mathcal{X} \setminus \mathcal{A}$  and  $\nu_i = 0$ ,  $\nu_j = \xi_j^* + \xi_i^*$ ,  $\nu_k = \xi_k^*$  for  $k \neq i, j$  on  $\mathcal{A}$ . We have

$$\begin{aligned} F_{H^{(\alpha)}}(\underline{\xi}^*; \underline{\nu}) &= (\mathbb{P}_{\nu_i} - \mathbb{P}_{\xi_i^*})G_i^{(\alpha)}(\underline{\xi}^*, \cdot) + (\mathbb{P}_{\nu_j} - \mathbb{P}_{\xi_j^*})G_j^{(\alpha)}(\underline{\xi}^*, \cdot) \\ &= \int_{\mathcal{A}} [G_j^{(\alpha)}(\underline{\xi}^*, x) - G_i^{(\alpha)}(\underline{\xi}^*, x)] d\xi_i^*(x) > 0 \end{aligned}$$

which implies that  $\xi^*$  cannot be optimal.

Proof that (ii)  $\implies$  (iii). For any subset  $\mathcal{J}_t$  of  $\{1, \dots, K\}$ , define

$$\begin{aligned} \mathcal{X}_{\mathcal{J}_t} &= \{x \in \mathcal{X} : G_i^{(\alpha)}(\underline{\xi}^*, x) = G_j^{(\alpha)}(\underline{\xi}^*, x) > G_k^{(\alpha)}(\underline{\xi}^*, x) \\ &\quad \text{for all } i, j \in \mathcal{J}_t \text{ and } k \notin \mathcal{J}_t\}. \end{aligned}$$

Let  $n_K$  denote the number of nonempty sets among those  $2^K - 1$  sets; they form a partition of  $\mathcal{X}$ . Take any  $\mathcal{X}_{\mathcal{J}_t}$  in this partition, we show that  $\sum_{i \in \mathcal{J}_t} \xi_i^* = \mu$  on  $\mathcal{X}_{\mathcal{J}_t}$  by contradiction. Indeed, suppose this is not true, then there exist a subset  $\mathcal{A}$  of  $\mathcal{X}_{\mathcal{J}_t}$  and  $k \notin \mathcal{J}_t$  such that  $\xi_k^*(\mathcal{A}) > 0$ ; (ii) implies that  $G_k^{(\alpha)}(\underline{\xi}^*, x) \geq G_i^{(\alpha)}(\underline{\xi}^*, x)$  for  $x \in \mathcal{A}$  and  $i \in \mathcal{J}_t$ , which contradicts the definition of  $\mathcal{X}_{\mathcal{J}_t}$ . The  $n_K$  nonempty sets  $\mathcal{X}_{\mathcal{J}_t}$  thus satisfy the required properties.

Proof that (iii)  $\implies$  (i). Let  $\underline{\nu}$  be element of  $\Xi(\mu)$ . We have

$$F_{H^{(\alpha)}}(\underline{\xi}^*; \underline{\nu}) = \sum_{t=1}^{n_K} \left\{ \sum_{k=1}^K \left( \int_{\mathcal{X}_{\mathcal{J}_t}} G_k^{(\alpha)}(\underline{\xi}^*, x) d\nu_k(x) - \int_{\mathcal{X}_{\mathcal{J}_t}} G_k^{(\alpha)}(\underline{\xi}^*, x) d\xi_k^*(x) \right) \right\}.$$

For each  $t$  and for all  $i \in \mathcal{J}_t$  and  $j \notin \mathcal{J}_t$ ,  $G_i^{(\alpha)}(\underline{\xi}^*, x) = G_{(t)}^{(\alpha)}(x) > G_j^{(\alpha)}(\underline{\xi}^*, x)$  on  $\mathcal{X}_{\mathcal{J}_t}$  and  $\xi_j^*(\mathcal{X}_{\mathcal{J}_t}) = 0$ . Therefore,

$$\begin{aligned} &\sum_{k=1}^K \left( \int_{\mathcal{X}_{\mathcal{J}_t}} G_k^{(\alpha)}(\underline{\xi}^*, x) d\nu_k(x) - \int_{\mathcal{X}_{\mathcal{J}_t}} G_k^{(\alpha)}(\underline{\xi}^*, x) d\xi_k^*(x) \right) \\ &\leq \int_{\mathcal{X}_{\mathcal{J}_t}} G_{(t)}^{(\alpha)}(x) \left( \sum_k d\nu_k(x) - \int_{\mathcal{X}_{\mathcal{J}_t}} G_{(t)}^{(\alpha)}(x) \left( \sum_{i \in \mathcal{J}_t} d\xi_i^*(x) \right) \right) = 0. \end{aligned}$$

We thus obtain that  $F_{H^{(\alpha)}}(\underline{\xi}^*; \underline{\nu}) \leq 0$  for any  $\underline{\nu} \in \Xi(\mu)$ , which implies that  $\underline{\xi}^*$  is optimal.  $\blacksquare$

*Proof of Theorem 2.2.* We denote  $\Delta_{12}(x) = G_1^{(\alpha)}(\underline{\xi}^*, x) - G_2^{(\alpha)}(\underline{\xi}^*, x)$ ,  $x \in \mathcal{X}$ .

Proof that (i)  $\implies$  (ii). Denote  $c_1 = \max\{c : \xi_1^*\{x \in \mathcal{X} : \Delta_{12}(x) \geq c\} = \xi_1^*(\mathcal{X})\}$ ,  $c_2 = \min\{c : \xi_2^*\{x \in \mathcal{X} : \Delta_{12}(x) \geq c\} = \xi_2^*(\mathcal{X})\}$ . Suppose that (ii) is not satisfied. Then,  $c_1 < c_2$  (otherwise,  $c = (c_1 + c_2)/2$  would satisfy (ii)). Define  $c_1^* = (2c_1 + c_2)/3$ ,  $c_2^* = (c_1 + 2c_2)/3$ . For  $\epsilon > 0$  small enough, we can always find two subsets  $\mathcal{A}_1$  and  $\mathcal{A}_2$  of  $\mathcal{X}$  such that  $\xi_1^*(\mathcal{A}_1) = \xi_2^*(\mathcal{A}_2) = \epsilon$  and  $\Delta_{12}(x) < c_1^*$  on  $\mathcal{A}_1$ ,  $\Delta_{12}(x) > c_2^*$  on  $\mathcal{A}_2$ . Consider the design  $\underline{\nu} = (\nu_1, \nu_2) \in \Xi(\mu)$ , defined by  $\underline{\nu} = \underline{\xi}$  on  $\mathcal{X} \setminus (\mathcal{A}_1 \cup \mathcal{A}_2)$  and

$$(\nu_1 = 0, \nu_2 = \xi_1^* + \xi_2^*) \text{ on } \mathcal{A}_1; (\nu_1 = \xi_1^* + \xi_2^*, \nu_2 = 0) \text{ on } \mathcal{A}_2.$$

It satisfies  $\nu_k(\mathcal{X}) = \xi_k^*(\mathcal{X})$ ,  $k = 1, 2$ , and therefore belongs to  $\Xi_\beta(\mu)$  (i.e., it satisfies  $\underline{\pi}(\underline{\nu}) \geq \beta$ ). The directional derivative  $F_{H^{(\alpha)}}(\underline{\xi}^*; \underline{\nu})$  satisfies

$$\begin{aligned} F_{H^{(\alpha)}}(\underline{\xi}^*; \underline{\nu}) &= \sum_{k=1}^2 [\mathbb{P}_{\nu_k} - \mathbb{P}_{\xi_k^*}] G_k^{(\alpha)}(\underline{\xi}^*, \cdot) \\ &= \int_{\mathcal{A}_2} \Delta_{12}(x) d\xi_2^*(x) - \int_{\mathcal{A}_1} \Delta_{12}(x) d\xi_1^*(x) \\ &> \epsilon(c_2^* - c_1^*) = \epsilon \frac{c_2 - c_1}{3} > 0, \end{aligned}$$

which contradicts the optimality of  $\underline{\xi}^*$ .

Proof that (ii)  $\implies$  (iii). We simply take  $\mathcal{X}_1 = \{x \in \mathcal{X} : \Delta_{12}(x) > c\}$  and  $\mathcal{X}_2 = \{x \in \mathcal{X} : \Delta_{12}(x) < c\}$ .

Proof that (iii)  $\implies$  (i). For any  $\underline{\nu} \in \Xi_\beta(\mu)$ , we have

$$\begin{aligned} F_{H^{(\alpha)}}(\underline{\xi}^*; \underline{\nu}) &= \int_{\mathcal{X}_1} \sum_{k=1}^2 G_k^{(\alpha)}(\underline{\xi}^*, x) (d\nu_k - d\xi_k^*)(x) \\ &+ \int_{\mathcal{X}_2} \sum_{k=1}^2 G_k^{(\alpha)}(\underline{\xi}^*, x) (d\nu_k - d\xi_k^*)(x) + \int_{\mathcal{X}_3} \sum_{k=1}^2 G_k^{(\alpha)}(\underline{\xi}^*, x) (d\nu_k - d\xi_k^*)(x), \end{aligned}$$

where we have denoted  $\mathcal{X}_3 = \mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$ . Since  $G_2^{(\alpha)}(\underline{\xi}^*, x) < G_1^{(\alpha)}(\underline{\xi}^*, x) - c$  on  $\mathcal{X}_1$ ,  $G_1^{(\alpha)}(\underline{\xi}^*, x) < G_2^{(\alpha)}(\underline{\xi}^*, x) + c$  on  $\mathcal{X}_2$ , and  $G_1^{(\alpha)}(\underline{\xi}^*, x) = G_2^{(\alpha)}(\underline{\xi}^*, x) + c$  on  $\mathcal{X}_3$ , and using the properties  $\nu_1 + \nu_2 = \xi_1^* + \xi_2^* = \mu$ ,  $\xi_1^* = \mu$  on  $\mathcal{X}_1$  and  $\xi_2^* = \mu$  on  $\mathcal{X}_2$ , we obtain

$$F_{H^{(\alpha)}}(\underline{\xi}^*; \underline{\nu}) \leq c[-\nu_2(\mathcal{X}_1) + \nu_1(\mathcal{X}_2) + \nu_1(\mathcal{X}_3) - \xi_1^*(\mathcal{X}_3)]. \quad (5.1)$$

Suppose that  $\underline{\xi}^*$  satisfies (iii) with  $c = 0$ . Then  $F_{H^{(\alpha)}}(\underline{\xi}^*; \underline{\nu}) \leq 0$  for all  $\underline{\nu} \in \Xi(\mu)$  and  $\underline{\xi}^*$  is optimal (this is in fact Corollary 2.1). Suppose now, without any loss of generality, that  $\underline{\xi}^*$  satisfies (iii) with  $c > 0$ . This means that the constraint

$\pi(\underline{\xi}^*) \geq \beta$  is saturated, and more precisely that  $\xi_2^*(\mathcal{X}) = \beta$ . Then, substituting  $\mu(\mathcal{X}_2) - \nu_2(\mathcal{X}_2)$  for  $\nu_1(\mathcal{X}_2)$ ,  $\mu(\mathcal{X}_3) - \nu_2(\mathcal{X}_3)$  for  $\nu_1(\mathcal{X}_3)$  and  $\mu(\mathcal{X}_3) - \xi_2^*(\mathcal{X}_3)$  for  $\xi_1^*(\mathcal{X}_3)$  in the upper bound (5.1), since  $\nu_2(\mathcal{X}) \geq \beta$  for  $\nu \in \Xi_\beta(\mu)$ , we obtain

$$F_{H^{(\alpha)}}(\underline{\xi}^*; \nu) \leq c[\xi_2^*(\mathcal{X}) - \nu_2(\mathcal{X})] \leq 0,$$

which implies that  $\underline{\xi}^*$  is optimal.  $\blacksquare$

In the proof of Theorem 4.1 we shall consider non-sequential exact designs  $\underline{\xi}_n^{\text{NS}} = (\xi_{n,1}^{\text{NS}}, \dots, \xi_{n,K}^{\text{NS}})$  that maximize  $H_n^{(\alpha)} = (1 - \alpha)\Psi(\mathbf{M}_n) + \alpha R_n$ , see (2.3). Next lemma indicates that  $\underline{\xi}_n^{\text{NS}}$  has asymptotically the same information and regret values as an optimal design  $\underline{\xi}^* \in \Xi(\mu)$ , see Section 2.

**Lemma 5.1** *Under H1a or (H1b, H2, H2'), a non-sequential exact design  $\underline{\xi}_n^{\text{NS}} = (\xi_{n,1}^{\text{NS}}, \dots, \xi_{n,K}^{\text{NS}})$  that maximizes  $H_n^{(\alpha)} = (1 - \alpha)\Psi(\mathbf{M}_n) + \alpha R_n$ ,  $\alpha \in [0, 1]$ , satisfies*

$$H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} H^{(\alpha)}(\underline{\xi}^*), \quad n \rightarrow \infty, \quad (5.2)$$

with  $\underline{\xi}^*$  an optimal design maximizing  $H^{(\alpha)}(\underline{\xi})$  with respect to  $\underline{\xi} \in \Xi(\mu)$ . Moreover,  $\psi(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} \psi(\underline{\xi}^*)$  if  $\alpha < 1$  and  $R(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} R(\underline{\xi}^*)$  if  $\alpha > 0$ .

*Proof of Lemma 5.1.* Any arbitrary sequential allocation rule satisfies  $H_n^{(\alpha)} = (1 - \alpha)\Psi(\mathbf{M}_n) + \alpha R_n \leq H^{(\alpha)}(\underline{\xi}_n^{\text{NS}})$ . This is true in particular for  $\underline{\xi}_n^*$  of Section 4.1. Since  $H^{(\alpha)}(\underline{\xi}_n^*) \xrightarrow{\text{a.s.}} H^{(\alpha)}(\underline{\xi}^*)$  as  $n \rightarrow \infty$ , we have  $\liminf_{n \rightarrow \infty} H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}) \geq H^{(\alpha)}(\underline{\xi}^*)$ . In the rest of the proof, we show that  $\limsup_{n \rightarrow \infty} H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}) \leq H^{(\alpha)}(\underline{\xi}^*)$  and then consider the limits of  $\psi(\underline{\xi}_n^{\text{NS}})$  and  $R(\underline{\xi}_n^{\text{NS}})$ .

Assume first that  $\alpha < 1$ . We treat the two cases H1a and (H1b, H2, H2') separately.

Under H1a, suppose that  $\mathcal{X} = \{x^{(1)}, \dots, x^{(m)}\}$  and, without any loss of generality, that  $\mu(x^{(j)}) > 0$  for all  $j$ . Denote  $\delta_n = [\max_{j=1, \dots, m} \mu_n(x^{(j)}) - \mu(x^{(j)})] / \mu(x^{(j)})$ , with  $\mu_n$  the empirical measure of the  $X_i$  for the first  $n$  subjects. From the SLLN,  $\delta_n \xrightarrow{\text{a.s.}} 0$ ,  $n \rightarrow \infty$ . Now,  $H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}) \leq H^{(\alpha)}(\underline{\xi}^*[\mu_n])$ , where  $\underline{\xi}^*[\mu_n]$  maximizes  $H^{(\alpha)}(\underline{\xi})$  with respect to  $\underline{\xi} \in \Xi(\mu_n)$  (since  $\underline{\xi}_n^{\text{NS}}$  is subject to the restrictions imposed to an exact design, whereas  $\underline{\xi}^*[\mu_n]$  is not). Without any loss of generality, we can assume that all  $\eta_k$  are positive (since optimal designs are invariant by addition of a positive constant to each  $\eta_k$ ). Then, since  $\mathbf{M}(\underline{\xi})$  is linear in  $\underline{\xi}$  and  $\psi(\cdot)$  is Loewner increasing, we have  $H^{(\alpha)}(\underline{\xi}^*[\mu_n]) \leq H^{(\alpha)}(\underline{\xi}^*[(1 + \delta_n)\mu])$ . From the concavity of  $H^{(\alpha)}(\cdot)$ , we finally obtain

$$\begin{aligned} H^{(\alpha)}(\underline{\xi}^*[(1 + \delta_n)\mu]) &\leq H^{(\alpha)}\left(\frac{1}{1 + \delta_n} \underline{\xi}^*[(1 + \delta_n)\mu]\right) \\ &+ (1 - \alpha) \text{trace} \left\{ \nabla \psi \left[ \frac{1}{1 + \delta_n} \mathbf{M}(\underline{\xi}^*[(1 + \delta_n)\mu]) \right] \frac{\delta_n}{1 + \delta_n} \mathbf{M}(\underline{\xi}^*[(1 + \delta_n)\mu]) \right\} \\ &+ \alpha \frac{\delta_n}{1 + \delta_n} \phi(\underline{\xi}^*[(1 + \delta_n)\mu]) \leq H^{(\alpha)}(\underline{\xi}^*[\mu]) + O(\delta_n), \end{aligned}$$

where the last inequality follows from the fact that  $\underline{\xi}^*[(1+\delta_n)\mu]/(1+\delta_n) \in \Xi(\mu)$ . Therefore,  $\limsup_{n \rightarrow \infty} H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}) \leq H^{(\alpha)}(\underline{\xi}^*)$ .

The proof under (H1b, H2, H2') is more involved, and we use results on empirical processes to avoid lengthy developments. For any design  $\underline{\xi}$  and any  $k$ , the function  $x \rightarrow G_k^{(\alpha)}(\underline{\xi}, x)$  depends on  $\underline{\xi}$  only through the matrix  $\mathbf{M}(\underline{\xi})$ , and this dependence is continuous in  $\mathbf{M}(\underline{\xi})$ . For any pair of designs  $\underline{\xi}_a, \underline{\xi}_b$ , we have

$$G_k^{(\alpha)}(\underline{\xi}_a, x) - G_k^{(\alpha)}(\underline{\xi}_b, x) = (1 - \alpha) \text{trace}\{[\nabla\psi(\underline{\xi}_a) - \nabla\psi(\underline{\xi}_b)]\mathfrak{M}_k(x)\}. \quad (5.3)$$

Since  $\alpha < 1$  and  $H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}) \geq H^{(\alpha)}(\underline{\xi}_n^*) \xrightarrow{\text{a.s.}} H^{(\alpha)}(\underline{\xi}^*)$ , there exists (a.s.) a  $n_1$  and  $\epsilon > 0$  such that  $\lambda_{\min}[\mathbf{M}(\underline{\xi}_n^{\text{NS}})] > \epsilon$  for all  $n > n_1$ , with  $\lambda_{\min}(\mathbf{M})$  the minimum eigenvalue of  $\mathbf{M}$ . Since  $\Psi(\cdot)$  is twice continuously differentiable in  $\mathbb{M}^+$ , (5.3) implies that  $|G_k^{(\alpha)}(\underline{\xi}_a, x) - G_k^{(\alpha)}(\underline{\xi}_b, x)| < B(\epsilon) \|\mathbf{M}(\underline{\xi}_a) - \mathbf{M}(\underline{\xi}_b)\| \sqrt{\text{trace}[\mathfrak{M}_k^2(x)]}$  when  $\lambda_{\min}[\mathbf{M}(\underline{\xi}_i)] > \epsilon$ ,  $i = a, b$ , for some  $B(\epsilon) < \infty$ . The functions  $x \rightarrow G_k^{(\alpha)}(\underline{\xi}_n^{\text{NS}}, x)$  thus belong to a Glivenko-Cantelli class (and even a Donsker class) for all  $n > n_1$  (a.s.), see, e.g., van der Vaart (1998, p. 271). With probability 1, the design  $\underline{\xi}_n^{\text{NS}}$  defines a partition of  $\mathcal{X}$  in  $K$  sets  $\mathcal{X}_{n,k}$  such that  $G_k^{(\alpha)}(\underline{\xi}_n^{\text{NS}}, x) = \max_{j=1, \dots, K} G_j^{(\alpha)}(\underline{\xi}_n^{\text{NS}}, x)$  for  $x \in \mathcal{X}_{n,k}$ , see Theorem 2.1. Denote by  $\underline{\xi}_n^{\text{NS}}[\mu] = (\underline{\xi}_{n,1}^{\text{NS}}[\mu], \dots, \underline{\xi}_{n,K}^{\text{NS}}[\mu]) \in \Xi(\mu)$  the design defined by  $\underline{\xi}_{n,k}^{\text{NS}}[\mu] = \mu$  for  $x \in \mathcal{X}_{n,k}$ . The concavity of  $\Psi(\cdot)$  gives

$$\begin{aligned} H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}) &\leq H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}[\mu]) + \sum_{k=1}^K \left[ \mathbb{P}_{\underline{\xi}_n^{\text{NS}}[\mu]} - \mathbb{P}_{\underline{\xi}_n^{\text{NS}}} \right] G_k^{(\alpha)}(\underline{\xi}_n^{\text{NS}}, \cdot) \\ &= H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}[\mu]) + [\mathbb{P}_\mu - \mathbb{P}_{\mu_n}] \max_k G_k^{(\alpha)}(\underline{\xi}_n^{\text{NS}}, \cdot) \\ &\xrightarrow{\text{a.s.}} H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}[\mu]) \leq H^{(\alpha)}(\underline{\xi}^*) \quad n \rightarrow \infty. \end{aligned}$$

This concludes the proof of (5.2) for  $\alpha < 1$ . Since  $\mathbf{M}(\underline{\xi}^*)$  is unique for  $\alpha < 1$ , see Section 2.3.2,  $\mathbf{M}(\underline{\xi}_n^{\text{NS}}[\mu]) \xrightarrow{\text{a.s.}} \mathbf{M}(\underline{\xi}^*)$  and  $\mathbf{M}(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} \mathbf{M}(\underline{\xi}^*)$ , so that  $\psi(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} \psi(\underline{\xi}^*)$ . Moreover,  $R(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} R(\underline{\xi}^*)$  if  $\alpha > 0$ .

When  $\alpha = 1$ , (5.2) and  $R(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} R(\underline{\xi}^*)$  follow from (2.1), since  $\underline{\xi}_n^{\text{NS}}$  coincides with  $\underline{\xi}_n^*$  of Section 4.1.  $\blacksquare$

When randomization is introduced, we can compare the performance of  $\underline{\xi}_n^{\text{NS}}$  with that obtained for  $\underline{\xi}_R^{*(\alpha, \beta)} \in \Xi(\mu)$  that maximizes  $H^{(\alpha)}(\underline{\xi})$  under the constraint  $r(\underline{\xi}) \geq \beta$ , see Section 3. The exact design  $\underline{\xi}_n^{\text{NS}}$  is now such that  $n_\beta = \lceil \beta n \rceil$  subjects are assigned randomly with uniform probability among treatments, the remaining  $n - n_\beta$  subjects being assigned optimally, in order to maximize  $H_n^{(\alpha)}$ . The same results as in Lemma 5.1 then apply, with the additional property that  $\mathbf{M}(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} \mathbf{M}(\underline{\xi}_R^{*(\alpha, \beta)})$  also for  $\alpha = 1$ , with  $\mathbf{M}(\underline{\xi}_R^{*(\alpha, \beta)})$  having full rank.

Since the functions  $x \rightarrow G_k^{(\alpha)}(\underline{\xi}_n^{\text{NS}}, x)$  belong to a Donsker class for  $n$  large

enough, we also have a CLT for  $\mathbf{M}(\underline{\xi}_n^{\text{NS}})$ ,  $\psi(\underline{\xi}_n^{\text{NS}})$  and  $R(\underline{\xi}_n^{\text{NS}})$ . However, the proof of Theorem 4.1 only uses the property  $H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} H^{(\alpha)}(\underline{\xi}_R^{*(\alpha,\beta)})$ .

*Proof of Theorem 4.1.* Assume first that  $\alpha < 1$ . For all  $n$ , denote  $H_n^{(\alpha)} = H^{(\alpha)}(\hat{\underline{\xi}}_n)$ , with  $\hat{\underline{\xi}}_n = (\hat{\xi}_{n,1}, \dots, \hat{\xi}_{n,K})$ , so that  $\hat{\underline{\xi}}_{n+1} = n/(n+1)\hat{\underline{\xi}}_n + 1/(n+1)\underline{\omega}_{n+1}$  with  $\omega_{n+1,k} = \bar{T}_{n+1,k}\delta X_{n+1}$  for all  $k$ . Using a second-order Taylor development, we can write

$$\begin{aligned} H_{n+1}^{(\alpha)} &= H_n^{(\alpha)} + \frac{1}{n+1} [G_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \underline{\omega}_{n+1}) - G_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \hat{\underline{\xi}}_n)] \\ &\quad + \frac{1-\alpha}{2(n+1)^2} \frac{\partial^2 \psi[(1-\gamma)\hat{\underline{\xi}}_n + \gamma\underline{\omega}_{n+1}]}{\partial \gamma^2} \Big|_{\gamma=\gamma_0} \end{aligned}$$

for some  $\gamma_0 \in (0, 1/(n+1))$ . Since  $\beta > 0$  and  $\mathbf{M}(\underline{\xi}_\mu) \in \mathbb{M}^+$  (with  $\xi_{\mu_k} = \mu/K$  for all  $k$ ), there exists  $\epsilon > 0$  and  $n_1$  (a.s.) such that for all  $n > n_1$  the minimum eigenvalue of  $\mathbf{M}[(1-\gamma_0)\hat{\underline{\xi}}_n + \gamma_0\underline{\omega}_{n+1}]$  is larger than  $\epsilon$ . Therefore,  $\partial^2 \psi[(1-\gamma)\hat{\underline{\xi}}_n + \gamma\underline{\omega}_{n+1}]/\partial \gamma^2|_{\gamma=\gamma_0} > -A(\epsilon)$  for some  $A(\epsilon) < \infty$ , and

$$H_{n+1}^{(\alpha)} > H_n^{(\alpha)} + \frac{1}{n+1} [G_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \underline{\omega}_{n+1}) - G_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \hat{\underline{\xi}}_n)] - \frac{(1-\alpha)A(\epsilon)}{2(n+1)^2}. \quad (5.4)$$

Since  $T_{n+1,k}$  is generated via (4.7), we obtain

$$\mathbb{E}\{G_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \underline{\omega}_{n+1}) - G_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \hat{\underline{\xi}}_n) | \mathcal{F}_n\} = \max_{\underline{\nu} \in \Xi(\mu), r(\underline{\nu}) \geq \beta} F_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \underline{\nu}),$$

and thus

$$\mathbb{E}\{H_{n+1}^{(\alpha)} | \mathcal{F}_n\} > H_n^{(\alpha)} + \frac{1}{n+1} \max_{\underline{\nu} \in \Xi(\mu), r(\underline{\nu}) \geq \beta} F_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \underline{\nu}) - \frac{(1-\alpha)A(\epsilon)}{2(n+1)^2}$$

for all  $n > n_1$ . Note that  $\max_{\underline{\nu} \in \Xi(\mu), r(\underline{\nu}) \geq \beta} F_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \underline{\nu})$  may be negative, so that a direct use of the Robbins-Siegmund's Theorem (1971) is not possible.

Since  $H_n \leq H^{(\alpha)}(\underline{\xi}_n^{\text{NS}}) \xrightarrow{\text{a.s.}} H_*^{(\alpha)} = H^{(\alpha)}(\underline{\xi}_R^{*(\alpha,\beta)})$ , see Lemma 5.1, we have  $\limsup_{n \rightarrow \infty} H_n \leq H_*^{(\alpha)}$ . Suppose that  $\limsup_{n \rightarrow \infty} H_n < H_*^{(\alpha)} - \delta$  for some  $\delta > 0$ . The concavity of  $H^{(\alpha)}(\cdot)$  implies that  $\max_{\underline{\nu} \in \Xi(\mu), r(\underline{\nu}) \geq \beta} F_{H^{(\alpha)}}(\hat{\underline{\xi}}_n, \underline{\nu}) > \delta/2$  for  $n$  large enough, and therefore  $\mathbb{E}\{H_{n+1}^{(\alpha)} | \mathcal{F}_n\} \geq H_n + \delta/[4(n+1)]$  for all  $n$  larger than some  $n_2$ , which is impossible (since  $H^{(\alpha)}(\underline{\xi})$  is bounded). This implies that  $\limsup_{n \rightarrow \infty} H_n = H_*^{(\alpha)}$  (a.s.), and we only need to show that, for any  $\delta > 0$ , the event  $\liminf_{n \rightarrow \infty} H_n < H_*^{(\alpha)} - \delta$  has probability zero. The result follows from arguments similar to those used for the proof of Doob's upcrossing Lemma, see Williams (1991, page 108): (5.4) implies that  $H_{n+1} - H_n > -\delta/6$  for all  $n$  large enough, while on the other hand  $\limsup_{n \rightarrow \infty} H_n = H_*^{(\alpha)}$  implies that  $H_n < H_*^{(\alpha)} + \delta/6$  for all  $n$  large enough. The rest of the proof is identical to Pronzato (2006, Th. 9).

Finally,  $\mathbf{M}(\hat{\xi}_n) \xrightarrow{\text{a.s.}} \mathbf{M}(\xi_R^{*(\alpha,\beta)})$  since the optimal matrix is unique, therefore  $\psi(\hat{\xi}_n) \xrightarrow{\text{a.s.}} \psi(\xi_R^{*(\alpha,\beta)})$ , and moreover  $R(\hat{\xi}_n) \xrightarrow{\text{a.s.}} R(\xi_R^{*(\alpha,\beta)})$  for  $\alpha > 0$ .

If  $\alpha = 1$ ,  $\hat{\xi}_n$  coincides with  $\xi_n^{\text{NS}}$  from (2.1), and Lemma 5.1 applies. Moreover,  $\mathbf{M}(\hat{\xi}_n) \xrightarrow{\text{a.s.}} \mathbf{M}(\xi_R^{*(1,\beta)})$  since  $\beta > 0$ . ■

## References

- Atkinson, A. (1982). Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika*, 69(1):61–67.
- Atkinson, A. (1999). Optimum biased-coin designs for sequential treatment allocation with covariate information. *Statistics in Medicine*, 18(14):1741–1752.
- Atkinson, A. (2002). The comparison of designs for sequential clinical trials with covariate information. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 165(2):349–373.
- Atkinson, A. and Biswas, A. (2005). Bayesian adaptive biased-coin designs for clinical trials with normal responses. *Biometrics*, 61:118–125.
- Baldi Antognini, A. and Giovagnoli, A. (2010). Compound optimal allocation for individual and collective ethics in binary clinical trials. *Biometrika*, 97(4):935–946.
- Baldi Antognini, A. and Zagoraiou, M. (2015). On the almost sure convergence of adaptive allocation procedures. *Bernoulli*, 21(2):881–908.
- Ball, F., Smith, A., and Verdinelli, I. (1993). Biased coin designs with a Bayesian bias. *Journal of Statistical Planning and Inference*, 34:403–421.
- Cappé, O., Garivier, A., Maillard, O.-A., Munos, R., and Stoltz, G. (2013). Kullback-Leibler upper confidence bounds for optimal sequential allocation. *Annals of Statistics*, 41(3):1516–1541.
- Cook, D. and Wong, W. (1994). On the equivalence between constrained and compound optimal designs. *Journal of the American Statistical Association*, 89(426):687–692.
- Dragalin, V. and Fedorov, V. (2006). Adaptive designs for dose-finding based on efficacy-toxicity response. *Journal of Statistical Planning and Inference*, 136:1800–1823.
- Dragalin, V., Fedorov, V., and Wu, Y. (2008a). Adaptive designs for selecting drug combinations based on efficacy-toxicity response. *Journal of Statistical Planning and Inference*, 138:352–373.
- Dragalin, V., Fedorov, V., and Wu, Y. (2008b). Two-stage design for dose-finding that accounts for both efficacy and safety. *Statistics in Medicine*, 27:5156–5176.
- Efron, B. (1971). Forcing a sequential experiment to be balanced. *Biometrika*, 58(3):403–417.
- Fedorov, V. (1989). Optimal design with bounded density: optimization algorithms of the exchange type. *Journal Statist. Planning and Inference*, 22:1–13.
- Fedorov, V. and Leonov, S. (2014). *Optimal Design for Nonlinear Response Models*. CRC Press, Boca Raton.

- Goldenshluger, A. and Zeevi, A. (2011). A note on performance limitations in bandit problems with side information. *IEEE Transactions on Information Theory*, 57(3):1707–1713.
- Goldenshluger, A. and Zeevi, A. (2013). A linear response bandit problem. *Stochastic Systems*, 3(1):230–261.
- Hu, F. and Rosenberger, W. (2006). *The Theory of Response-Adaptive Randomization in Clinical Trials*. John Wiley & Sons.
- Hu, F. and Zhang, L.-X. (2004). Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. *Annals of Statistics*, 32(1):268–301.
- Hu, J., Zhu, H., and Hu, F. (2015). A unified family of covariate-adjusted response-adaptive designs based on efficiency and ethics. *Journal of the American Statistical Association*, 110(509):357–367.
- Hu, Y. and Hu, F. (2012). Asymptotic properties of covariate-adaptive randomization. *Annals of Statistics*, 40(3):1794–1815.
- Lai, T. and Robbins, H. (1985). Asymptotically efficient adaptive allocation rules. *Advances in Applied Math.*, 6:4–22.
- Pocock, S. and Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*, 31(1):103–115.
- Pronzato, L. (2004). A minimax equivalence theorem for optimum bounded design measures. *Statistics & Probability Letters*, 68:325–331.
- Pronzato, L. (2006). On the sequential construction of optimum bounded designs. *Journal of Statistical Planning and Inference*, 136:2783–2804.
- Pronzato, L. (2010). Penalized optimal designs for dose-finding. *Journal of Statistical Planning and Inference*, 140:283–296.
- Pronzato, L. and Pázman, A. (2013). *Design of Experiments in Nonlinear Models. Asymptotic Normality, Optimality Criteria and Small-Sample Properties*. Springer, LNS 212, New York, Heidelberg.
- Pshenichnyi, B. (1971). *Necessary Conditions for an Extremum*. Marcel Dekker, New York.
- Pukelsheim, F. (1993). *Optimal Experimental Design*. Wiley, New York.
- Rabie, H. and Flournoy, N. (2013). Optimal designs for contingent response models with application to toxicity-efficacy studies. *Journal of Statistical Planning and Inference*, 143:1371–1379.
- Robbins, H. and Siegmund, D. (1971). A convergence theorem for non negative almost supermartingales and some applications. In Rustagi, J., editor, *Optimization Methods in Statistics*, pages 233–257. Academic Press, New York.
- Rosenberger, W. and Lachin, J. (2002). *Randomization in Clinical Trials. Theory and Practice*. Wiley, New York.
- Sahm, M. and Schwabe, R. (2001). A note on optimal bounded designs. In Atkinson, A., Bogacka, B., and Zhigljavsky, A., editors, *Optimum Design 2000*, chapter 13, pages 131–140. Kluwer, Dordrecht.
- van der Vaart, A. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.
- Williams, D. (1991). *Probability with Martingales*. Cambridge University Press,

Cambridge.

- Wu, C. and Wynn, H. (1978). The convergence of general step-length algorithms for regular optimum design criteria. *Annals of Statistics*, 6(6):1273–1285.
- Wynn, H. (1982). Optimum submeasures with applications to finite population sampling. In Gupta, S. and Berger, J., editors, *Statistical Decision Theory and Related Topics III. Proc. 3rd Purdue Symp., vol. 2*, pages 485–495. Academic Press, New York.
- Zhang, L.-X. and Hu, F. (2009). A new family of covariate-adjusted response-adaptive designs and their properties. *Appl. Meth. J. Chinese Univ.*, 24(1):1–13.
- Zhang, L.-X., Hu, F., Cheung, S., and Chan, W. (2007). Asymptotic properties of covariate-adjusted response adaptive designs. *Annals of Statistics*, 35(3):1166–1182.