Information-regret compromise in covariate-adaptive treatment allocation

Asya Metelkina, Luc Pronzato

The Annals of Statistics
2017, Vol. 45, No. 5, 2046–2073
DOI: 10.1214/16-AOS1518

INFORMATION-REGRET COMPROMISE IN COVARIATE-ADAPTIVE TREATMENT ALLOCATION

BY ASYA METELKINA AND LUC PRONZATO

Université Côte d'Azur, CNRS, I3S

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 μ, its solution possesses some similarities with the construction of optimal design measures bounded by μ. 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.

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; see, for example, [22, 24, 39]. The response $Y_t = Y_t(X)$ of a subject with covariates X to treatment t satisfies

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

where θ_k denotes the (unknown) vector of model parameters for treatment k and where the functions η_k are assumed to be 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 responses η_k, but the multivariate situation case may be considered as well; see, for example, [13–15, 31] for bivariate binary responses corresponding to efficacy and toxicity. Note that the different models may have some parameters in common; that is, the vectors θ_k may share some components. We suppose that the covariates are i.i.d. among subjects, with some probability measure μ. The responses are independent too; that is, the random vectors $(X_i, Y_1(X_i), \dots, Y_k(X_i))$ are i.i.d.

Received February 2016; revised September 2016.
MSC2010 subject classifications. Primary 62K05; secondary 62P10.
Key words and phrases. Optimal design, treatment allocation, bounded design measure, equivalence theorem, multi-armed bandit problem.

Most clinical trials are designed in such a way that patients are assigned sequentially with some randomized rule. Their covariate information can then be used to tune allocation probabilities: in a typical Covariate-Adjusted Response-Adaptive (CARA) design, the $(n+1)$st patient, with covariates X_{n+1}, is assigned to treatment k with probability $\pi_k(X_{n+1}, \hat{\theta}^n)$, where $\hat{\theta}^n$ denotes the current estimated value of the vector θ of all parameters θ_k in the K models; see [21], Chapter 9, [7], Chapter 6, [39]. Extension to additional dependence in the past covariate values is considered in [8, 38]. The survey [33] contains a classification of CARA designs and the recent book [7] offers a thorough review of results of adaptive design theory. A key issue here, addressed in particular in [8, 38, 39], concerns the investigation of the asymptotic properties of $\hat{\theta}^n$ and allocation proportions. On the other hand, the definition of suitable desired allocation probabilities $\pi_k^*(x, \theta)$, in connection with some criterion for optimal design, is considered as an open problem, which forms the main motivation for this paper. Once suitable targets $\pi_k^*(x, \theta)$ are defined, results presented in [8, 24, 38, 39] can be used to construct procedures ensuring fast convergence to the $\pi_k^*(x, \theta)$ with low variability.

Clinical trials are usually facing two conflicting objectives: (i) statistical inference about the response models η_k; (ii) individual ethics, related to the efficient treatment of the individuals enroled in the trial. Here, these conflicting objectives will be taken into account *explicitly*, through (i) an information criterion related to the precision of θ and (ii) a cumulative regret relative to allocation of each subject to the best treatment. The overall strategy may then correspond to maximizing information under a constraint on regret, or minimizing 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 treatments of individuals in the trial, in the same spirit as what is done in [13, 14, 29] for dose selection (without covariates).

The main objective of the paper is to characterize optimal compromise designs and to derive simple covariate-adaptive rules $\pi_k^*(x, \theta)$ whose information and regret attain asymptotically the optimal values for the chosen compromise. The concept of design measures is natural for investigating asymptotic properties of designs, and we shall decompose μ into positive measures $\xi_1, \dots, \xi_K$ on the set of covariates, with ξ_k the fraction of μ corresponding to subjects allocated to treatment k. Any such decomposition defines a design $\underline{\xi}$. Due to the presence of the constraint $\sum_{k=1}^K \xi_k = \mu$ on $\underline{\xi}$, the optimal design problem presents some similarities with the construction of optimal bounded design measures (see [17, 34, 37]), and an equivalence theorem characterizing optimal designs $\underline{\xi}^*[\mu]$ is derived in Section 2. Although we do not explicitly aim at reducing imbalance between treatments, the avoidance of strongly imbalanced allocation is considered in Section 2.6. In biased-coin designs, a compromise has to be done between loss (of information concerning the estimation of treatment difference) and reduction of selection bias; see, for example, [5, 9, 24]; see also the recent survey [4] and the references therein. Our approach is different, since we target a compromise between information and allocation to the best treatment. As a consequence, our "optimal" allocation rules are deterministic. However, the introduction of randomization to reduce selection bias is considered in Section 3. 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 μ. The paper focusses on locally optimum design, where the model parameters θ_k are fixed to some prior nominal values. CARA rules, where allocation of the $(n+1)$st subject to one of the K treatments depends on the current X_{n+1} and estimated value $\hat{\theta}^n$, are briefly considered in Section 5, where other possible extensions are also suggested.

2. Optimal covariate-adaptive design.

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

H1a: $\mathscr{X}$ is finite;
H1b: $\mathscr{X}$ is a compact subset of $\mathbb{R}^d$ with nonempty interior $\text{int}(\mathscr{X})$ and μ has a density with respect to the Lebesgue measure.

In the second case, we shall assume the following:

H2: in (1.1), the functions $\eta_k(x, \theta_k)$ are continuously differentiable with respect to $x \in \text{int}(\mathscr{X})$.

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

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

Regret. If the parameters θ_k in each model η_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 θ_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 ith subject with covariates X_i has been allocated to treatment k_i, for all $i = 1, \dots, n$. This can also be written as

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

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

$$\xi_{n,k} = \frac{N_{n,k}}{n}\mu_{n,k}, \qquad k = 1, \dots, K, \tag{2.3}$$

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 η_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 θ, 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 jth row and column are formed of zeros when θ_k does not contain the jth component of θ. For example, in Bernoulli trials with a single response $Y_t \in \{0, 1\}$, we have

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

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

H2′: the functions $\{\mathfrak{M}_k(x; \theta_k)\}_{ij}$ are continuously differentiable with respect to $x \in \mathrm{int}(\mathscr{X})$, $1 \le i, j \le p$.

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

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

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 nonnegative 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

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

with A-optimal design when $q = 1$; see, for example, [30], Chapter 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 μ 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 $\mu_n \to \mu$ as $n \to \infty$ since the X_i are i.i.d. with μ. We shall thus consider design measures $\underline{\xi} = (\xi_1, \dots, \xi_K)$ that form a decomposition of μ into $\mu = \sum_{k=1}^{K} \xi_k$, where ξ_k will define the target limiting allocation in a sequential allocation rule and corresponds to the fraction of μ devoted to treatment k. Note that the ξ_k are not probability measures. We shall denote by

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

the (convex) set of such $\underline{\xi}$, with $\mathscr{M}_{\mathscr{X}}$ the set of nonnegative measures on $\mathscr{X}$ absolutely continuous with respect to μ.

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

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

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

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

see (2.5). We suppose that $\mathbf{M}(\underline{\xi}; \theta)$ is positive definite when all $\xi_k = \mu/K$ in $\underline{\xi}$; this balanced design will be denoted $\underline{\xi}_{\mu}$. 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 $\alpha \in [0, 1]$. The use of compromise designs in clinical trials is motivated in [6], but without considering the presence of covariates. Balancing efficiency and ethics is also considered in [23], but no explicit optimal design is used as a target for sequential allocation (see, in particular, Example 3 below). An alternative approach is considered in [10] for longitudinal responses with two treatments, where, for each X_{n+1}, the allocation probability $\pi_k(X_{n+1})$ is obtained by maximizing an utility function (depending on the current estimated value $\hat{\theta}^n$ and associated information matrix) that sets a compromise between information and allocation to the best treatment. Simpler sequential allocation rules targeting an optimal compromise will be proposed in Section 4.

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. Due to the equivalence between constrained and compound optimal designs (see [12, 29] and [18], Chapter 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 τ. Indeed, the Lagrangian for this constrained problem can be written as $\mathscr{L}(\underline{\xi}, C) = \psi(\underline{\xi};\theta) - C[R(\underline{\xi};\theta) - \tau]$, with $C \geq 0$ the Lagrange parameter; maximizing $\psi(\underline{\xi};\theta)$ with $R(\underline{\xi};\theta) \leq \tau$ is equivalent to maximizing $\mathscr{L}(\underline{\xi}, C)$ 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 nondominated solution $\underline{\xi}^*$ [i.e., 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 τ'.

We shall omit the dependence in θ 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

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

where $\phi(\underline{\xi}) = \sum_{k=1}^{K} \mathsf{P}_{\xi_k}\eta_k(\cdot)$ is a cumulative reward (to be maximized). The criterion (2.9) will be our measure of optimality in all the rest of the paper. Note that when $\Psi(\mathbf{M}) = \log\det(\mathbf{M})$ (D-optimality), optimal designs for $H^{(\alpha)}(\cdot)$ are invariant by reparameterization of the models η_k for any $\alpha \in [0, 1]$.

2.2. *An equivalence theorem for compromise optimal designs*. When μ is known, the maximization of $H^{(\alpha)}(\underline{\xi})$ under H1a corresponds to a finite dimensional concave problem. Indeed, for $\mathscr{X} = \{x^{(1)}, \dots, x^{(m)}\}$, we need to determine $k \times m$ weights $w_{k,j} = \xi_k(x^{(j)})$ satisfying $\sum_{k=1}^{K} w_{k,j} = \mu(x^{(j)})$ for all $j = 1, \dots, m$ and such that $H^{(\alpha)}(\underline{\xi})$ is maximum. The necessary and sufficient condition for optimality presented below provides a characterization of optimal designs that facilitates their construction in more general situations. The fact that optimal designs have particularly simple shapes will also facilitate the sequential constructions of Section 4.

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 $F_{H^{(\alpha)}}(\underline{\xi}; \underline{\nu})$ denotes the directional derivative:

$$F_{H^{(\alpha)}}(\underline{\xi}; \underline{\nu}) = \lim_{\gamma \to 0^+} \frac{H^{(\alpha)}((1-\gamma)\underline{\xi} + \gamma\underline{\nu}) - H^{(\alpha)}(\underline{\xi})}{\gamma}.$$

For any $\underline{\xi} \in (\mathscr{M}_{\mathscr{X}})^K$ such that $\Psi(\cdot)$ is differentiable at $\mathbf{M}(\underline{\xi})$, we denote $\nabla\psi(\underline{\xi}) = \nabla\Psi[\mathbf{M}(\underline{\xi})]$, where $\nabla\Psi(\mathbf{M}) = \partial\Psi(\mathbf{M})/\partial\mathbf{M}$. Direct calculation then gives

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

Denoting

$$G_k^{(\alpha)}(\underline{\xi}, x) = (1-\alpha)\operatorname{tr}\big[\nabla\psi(\underline{\xi})\mathfrak{M}_k(x)\big] + \alpha\eta_k(x), \tag{2.10}$$

we thus have $F_{H^{(\alpha)}}(\underline{\xi}; \underline{\nu}) = \sum_{k=1}^K \mathsf{P}_{\nu_k} G_k^{(\alpha)}(\underline{\xi}, \cdot) - \sum_{k=1}^K \mathsf{P}_{\xi_k} G_k^{(\alpha)}(\underline{\xi}, \cdot)$. We then obtain the following characterization of optimal design measures.

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, that is, $\underline{\xi}^*$ maximizes $H^{(\alpha)}(\underline{\xi})$ with respect to $\underline{\xi} \in \Xi(\mu)$*;

(ii) *for all $i = 1, \dots, K$, $G_i^{(\alpha)}(\underline{\xi}^*, x) \geq \max_{j\neq i} G_j^{(\alpha)}(\underline{\xi}^*, x)$ ξ_i^*-a.e.*;

(iii) *$\mathscr{X}$ can be partitioned into n_K subsets $\mathscr{X}_{\mathscr{J}_t} = \mathscr{X}_{\mathscr{J}_t}^{(\alpha)}$, $t = 1, \dots, n_K \leq 2^K - 1$, with index sets $\mathscr{J}_t$ that are subsets of $\{1, \dots, K\}$, and such that, for all $t = 1, \dots, n_K$*,

   (a) *$\sum_{i\in\mathscr{J}_t} \xi_i^* = \mu$ on $\mathscr{X}_{\mathscr{J}_t}$*,

   (b) *$G_i^{(\alpha)}(\underline{\xi}^*, x) = G_j^{(\alpha)}(\underline{\xi}^*, x)$ on $\mathscr{X}_{\mathscr{J}_t}$ for all $i, j \in \mathscr{J}_t$*,

   (c) *$G_i^{(\alpha)}(\underline{\xi}^*, x) > G_j^{(\alpha)}(\underline{\xi}^*, x)$ for $x \in \mathscr{X}_{\mathscr{J}_t}$, $i \in \mathscr{J}_t$ and $j \notin \mathscr{J}_t$.*

The proof of Theorem 2.1 is presented in the supplemental material [26]. The theorem takes a simpler form when $K = 2$: the function $x \mapsto \Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) = G_1^{(\alpha)}(\underline{\xi}^*, x) - G_2^{(\alpha)}(\underline{\xi}^*, x)$ then defines a partition of $\mathscr{X}$ 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, that is, it maximizes $H^{(\alpha)}(\underline{\xi})$ with respect to $\underline{\xi} \in \Xi(\mu)$*;

(ii) *$\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) \geq 0$ ξ_1^*-a.e. and $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) \leq 0$ ξ_2^*-a.e.*;

(iii) *there exist two subsets $\mathscr{X}_1 = \mathscr{X}_1^{(\alpha)}$ and $\mathscr{X}_2 = \mathscr{X}_2^{(\alpha)}$ of $\mathscr{X}$ such that*:

   (a) *$\xi_1^* = \mu$ on $\mathscr{X}_1$ and $\xi_2^* = \mu$ on $\mathscr{X}_2$*,

   (b) *$\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) = 0$ on $\mathscr{X} \setminus (\mathscr{X}_1 \cup \mathscr{X}_2)$*,

   (c) *$\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) > 0$ for $x \in \mathscr{X}_1$ and $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) < 0$ for $x \in \mathscr{X}_2$.*

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

2.3. *Some properties of 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 α 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 α in $[0,1]$, that $\psi(\underline{\xi}^{*(\alpha)})$ is nonincreasing for $\alpha \in [0,1)$ and $\phi(\underline{\xi}^{*(\alpha)})$ is nondecreasing on $(0,1]$, and that $H^{(\alpha)}(\underline{\xi}^{*(\alpha)})$ is convex and continuously differentiable with respect to α in $(0,1)$, with $\mathrm{d}H^{(\alpha)}(\underline{\xi}^{*(\alpha)})/\mathrm{d}\alpha = \phi(\underline{\xi}^{*(\alpha)}) - \psi(\underline{\xi}^{*(\alpha)})$.

2.3.1. *The special case* $\alpha = 1$. It corresponds to the usual framework in bandit theory, with abundant results on the construction of strategies minimizing the expected regret; see, for example, [19, 20, 25]. Theorem 2.1 applies when $\alpha = 1$ since $H^{(1)}(\underline{\xi}) = \sum_{k=1}^{K} \mathsf{P}_{\xi_k}\eta_k(\cdot)$ is linear in the ξ_k, and the sets $\mathscr{X}_j$ 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 $\mathscr{X}_1^{(1)} = \{x \in \mathscr{X} : \eta_1(x) > \eta_2(x)\}$ and $\xi_2^* = \mu$ on $\mathscr{X}_2^{(1)} = \{x \in \mathscr{X} : \eta_1(x) < \eta_2(x)\}$. Therefore, when the models are such that $\eta_1(x) > \eta_2(x)$ for all $x \in \mathscr{X}$, $\mathscr{X}_1^{(1)} = \mathscr{X}$ and $\underline{\xi}^*$ does not allow the estimation of θ_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 θ; 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 α in some interval $[\underline{\alpha}, 1]$; see Example 3. Note that

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

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 $\Psi(\mathbf{M}) = (\mathrm{tr}[\mathbf{M}^{-q}]/p)^{-1/q}$ $(q > 0)$, which have continuous extensions $\Psi(\mathbf{M}) = 0$ at singular $\mathbf{M}$; see [30], Chapter 5. Indeed, the property that the directional derivative of $\Psi(\cdot)$ at $\mathbf{M}$ in the direction $\mathbf{M}'$ 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 $x \mapsto G_k^{(\alpha)}(\underline{\xi}^{*(\alpha)}, x)$ and sets $\mathscr{X}_{\mathscr{J}_t}^{(\alpha)}$ 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

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

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

Condition (2.12) is often satisfied when $\alpha \in [0, 1)$ and μ 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 $\alpha = 0$, $K = 2$, η_1, η_2 have no parameters in common, but the numerical values of θ_1 and θ_2 are such that, for any ξ, $\mathbf{M}_1(\xi) = c\mathbf{M}_2(\xi)$ when we write $\mathbf{M}(\underline{\xi})$ as

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

with c some positive constant. Take for instance $\Psi(\mathbf{M}) = \log\det(\mathbf{M})$. Then, for any $\underline{\xi} \in \Xi(\mu)$ we have $H^{(0)}(\underline{\xi}) = \log\det[\mathbf{M}_1(\xi_1)] + \log\det[\mathbf{M}_1(\xi_2)] + p_1 \log(c)$, with p_1 the number of parameters in η_1 (and η_2). Consider $\underline{\xi}_\mu = (\mu/2, \mu/2) = ((\xi_1 + \xi_2)/2, (\xi_1 + \xi_2)/2) \in \Xi(\mu)$; it satisfies $H^{(0)}(\underline{\xi}_\mu) = 2\log\det\{\mathbf{M}_1[(\xi_1 + \xi_2)/2]\} + p_1 \log(c)$. The concavity of $\log\det(\cdot)$ implies that $H^{(0)}(\underline{\xi}_\mu) \geq H^{(0)}(\underline{\xi})$, and $\underline{\xi}_\mu$ is thus optimal, with $\mathscr{X}_1 = \mathscr{X}_2 = \varnothing$ in Corollary 2.1(iii). Moreover, any optimal design $\underline{\xi}^* = (\xi_1^*, \xi_2^*)$ is such that $\mathbf{M}_1(\xi_1^*) = \mathbf{M}_1(\xi_2^*)$, and the designs (ξ_2^*, ξ_1^*) and $(1-\gamma)\underline{\xi}^* + \gamma\underline{\xi}_\mu$ are also optimal for all $\gamma \in [0, 1]$. Since $\alpha = 0$, these optimal designs may have different regret values; see Example 3.

Provided that (2.12) is satisfied, so that $\underline{\xi}^{*(\alpha)}$ 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 $\xi_{n,k}(\mathscr{X})$ converge a.s. to their optimal counterparts $\xi_k^{*(\alpha)}(\mathscr{X})$. More generally, for $\alpha < 1$ this convergence of allocation proportions can always be ensured by including an intercept θ_{0k} in each of the k models η_k: indeed, for any $\underline{\xi}$ the diagonal element of $\mathbf{M}(\underline{\xi})$ corresponding to θ_{0k} equals $\xi_k(\mathscr{X})$, and the uniqueness of $\mathbf{M}(\underline{\xi}^{*(\alpha)})$ (see Section 2.3.2) implies that $\xi_{n,k}(\mathscr{X})$ converges to $\xi_k^{*(\alpha)}(\mathscr{X})$.

2.4. *Bounds on optimal regret and information.* The construction of an upper bound on the optimal regret and of a lower bound on the optimal information may help to choose a suitable compromise parameter α; see also Section 5.

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

$$\eta_j(x) - \eta_k(x) \leq \frac{1-\alpha}{\alpha} \operatorname{tr}\{\nabla\psi(\underline{\xi}^{*(\alpha)})[\mathfrak{M}_k(x) - \mathfrak{M}_j(x)]\};$$

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

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

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

$$R(\underline{\xi}^{*(\alpha)}) \leq \frac{1-\alpha}{\alpha} \operatorname{tr}[\nabla\psi(\underline{\xi}^{*(\alpha)})\mathbf{M}(\underline{\xi}^{*(\alpha)})]. \tag{2.13}$$

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 $R(\underline{\xi}^{*(\alpha)}) \leq p(1-\alpha)/\alpha$.

The optimality of $\underline{\xi}^{*(\alpha)}$ for $H^{(\alpha)}(\cdot)$ directly implies an upper bound on $R(\underline{\xi}^{*(\alpha)})$. Indeed, for any $\xi \in \Xi(\mu)$ we have

$$(1-\alpha)\psi(\underline{\xi}^{*(\alpha)}) + \alpha\phi(\underline{\xi}^{*(\alpha)}) \geq (1-\alpha)\psi(\underline{\xi}) + \alpha\phi(\underline{\xi}), \tag{2.14}$$

which, using (2.11), implies $R(\underline{\xi}^{*(\alpha)}) \leq R(\underline{\xi}) + (1-\alpha)/\alpha[\psi(\underline{\xi}^{*(\alpha)}) - \psi(\underline{\xi})]$ for $\alpha > 0$. Since $\psi(\underline{\xi}^{*(\alpha)}) \leq \psi(\underline{\xi}^{*(0)})$, we get

$$R(\underline{\xi}^{*(\alpha)}) \leq R(\underline{\xi}) + \frac{1-\alpha}{\alpha}[\psi(\underline{\xi}^{*(0)}) - \psi(\underline{\xi})], \qquad \forall \underline{\xi} \in \Xi(\mu), \alpha \in (0, 1]. \tag{2.15}$$

A lower bound on optimal information. Using (2.14) and (2.11), we obtain

$$\psi(\underline{\xi}^{*(\alpha)}) \geq \psi(\underline{\xi}) - \frac{\alpha}{1-\alpha} R(\underline{\xi}), \qquad \forall \underline{\xi} \in \Xi(\mu), \alpha \in [0, 1). \tag{2.16}$$

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 μ is uniform on $\mathscr{X} = [-1, 1]$ and that $Y_t|X = x, t = k$ has mean $\eta_k(x, \theta_k)$ and variance 1. We thus have, for any $x \in \mathscr{X}$ and any $\underline{\xi} = (\xi_1, \xi_2) \in \Xi(\mu)$,

$$\mathfrak{M}_1(x) = \begin{pmatrix} 1 & x & 0 \\ x & x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathfrak{M}_2(x) = \begin{pmatrix} 1 & 0 & x \\ 0 & 0 & 0 \\ x & 0 & x^2 \end{pmatrix},$$

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

where $m_k = \int_{\mathscr{X}} x \, \mathrm{d}\xi_k(x)$, $s_k^2 = \int_{\mathscr{X}} x^2 \, \mathrm{d}\xi_k(x)$ (with $m_1 + m_2 = 0$ and $s_1^2 + s_2^2 = \int_{\mathscr{X}} x^2 \, \mathrm{d}\mu(x) = 1/3$). This example illustrates the fact that we can have $\mathscr{X}_1^{(\alpha)} = \mathscr{X}_2^{(\alpha)} = \varnothing$ in Corollary 2.1 for all α in some interval also in the case where μ 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 $\mathscr{X}_1^{(\alpha)} = [-1, 0)$ and $\mathscr{X}_2^{(\alpha)} = (0, 1]$.

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 $x \in \mathscr{X}$ by choosing a $\underline{\xi} \in \Xi(\mu)$ with suitable values of m_1, m_2, s_1^2 and s_2^2. The conditions are $s_1^2 = s_2^2 = 1/6$ and $m_1 = -m_2 = m_1^{*(\alpha)}$ with

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

and then $\mathscr{X}_1^{(\alpha)} = \mathscr{X}_2^{(\alpha)} = \varnothing$ in Corollary 2.1. 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 $\lim_{\alpha \to 0} A(\alpha) = 1/\sqrt{2}$. Also, elementary calculations show that the optimal solution whose (Shannon) entropy $\sum_{i=1}^2 \int_{\mathscr{X}} -\log\{[\mathrm{d}\xi_i^*/\mathrm{d}\mu](x)\} \, \mathrm{d}\xi_i^*(x)$ is maximum is given by $[\mathrm{d}\xi_1^*/\mathrm{d}\mu](x) = (1/2)[1 + \exp(\lambda_1 x)]^{-1}$, with $\lambda_1 > 0$ such that $m_1 = m_1^{*(\alpha)}$. Note that $\xi_1^{*(\alpha)}(\mathscr{X}) = \xi_2^{*(\alpha)}(\mathscr{X}) = 1/2$ for all α for these two types of solutions, but other optimal designs $\underline{\xi}^{*(\alpha)}$ exist such that $\xi_1^{*(\alpha)}(\mathscr{X}) \neq \xi_2^{*(\alpha)}(\mathscr{X})$ 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

$$\psi(\underline{\xi}^{*(\alpha)}) = -2\log(12), \qquad R(\underline{\xi}^{*(\alpha)}) = 0 \qquad \text{for } \alpha \geq \underline{\alpha}$$

$$\psi(\underline{\xi}^{*(\alpha)}) = \log\left[\frac{1}{36} - \frac{(m_1^{*(\alpha)})^2}{3} \right],$$

$$R(\underline{\xi}^{*(\alpha)}) = (b_2 - b_1)\left[m_1^{*(\alpha)} + \frac{1}{4} \right] \qquad \text{otherwise.}$$

2.5.2. *Example* 2. This 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, $\theta = (a_0, b_2)^\top$; μ is still uniform on $\mathscr{X} = [-1, 1]$, $\Psi(\mathbf{M}) = \log\det(\mathbf{M})$. 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 $m_2 = \int_{\mathscr{X}} x \, d\xi_2(x)$, $s_2^2 = \int_{\mathscr{X}} x^2 \, d\xi_2(x)$. The design $\underline{\xi}^{*(1)}$ with $\mathscr{X}_1^{(\alpha)} = [-1, 0)$ and $\mathscr{X}_2^{(\alpha)} = (0, 1]$ 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 $\mathscr{X}_1^{(\alpha)} = (A(\alpha), 0)$ and $\xi_2^{*(\alpha)} = \mu$ on $\mathscr{X}_2^{(\alpha)} = [-1, A(\alpha)) \cup (0, 1]$, 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 $\alpha = 0$, observing Y_2 only permits to estimate both a_0 and b_2, while the 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 $\eta_1(x) = \eta_1(x, \theta_1) = 0.25 + 0.5e^{z_1(x)}/(1 + e^{z_1(x)})$, $\eta_2(x) = \eta_2(x, \theta_2) = 0.25 + 0.5/(1 + e^{z_2(x)})$, 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 μ uniform on $\mathscr{X} = [0, 1]$. 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 (and symmetric) on purpose.

For any $\alpha \in (0, 1]$, the optimal design $\underline{\xi}^* = \underline{\xi}^{*(\alpha)}$ is uniquely defined, with $\xi_1^* = \mu$ on $\mathscr{X}_1^{(\alpha)} = (A(\alpha), B(\alpha)) \cup (C(\alpha), 1]$ and $\xi_2^* = \mu$ on $\mathscr{X}_2^{(\alpha)} = [0, A(\alpha)) \cup (B(\alpha), C(\alpha))$ 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 α varying in $(0, 1]$

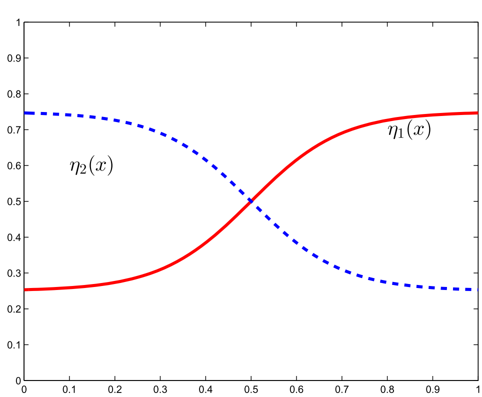

FIG. 1. *Example* 3: $\eta_1(x)$ (*solid line*) *and* $\eta_2(x)$ (*dashed line*) *as functions of* x.

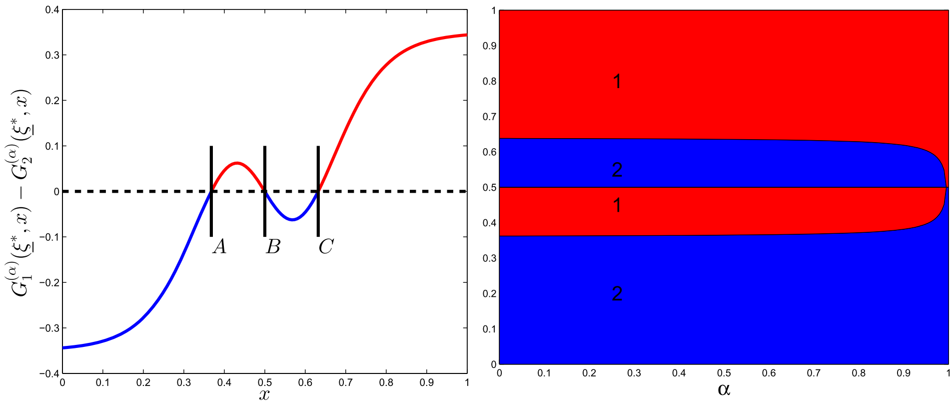

FIG. 2. *Example* 3—*Left*: *illustration of the optimality of* $\underline{\xi}^*$ *for* $\alpha = 0.7$; *right*: *subsets* $\mathscr{X}_1{}^{(\alpha)}$ *and* $\mathscr{X}_2{}^{(\alpha)}$ *of Corollary* 2.1(iii) *for* $\alpha \in (0, 1]$.

are shown in Figure 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 α. 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 α; the bottom part of the figure shows $\psi(\underline{\xi}^{*(\alpha)})$ (solid line) and the lower bound (2.16) (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

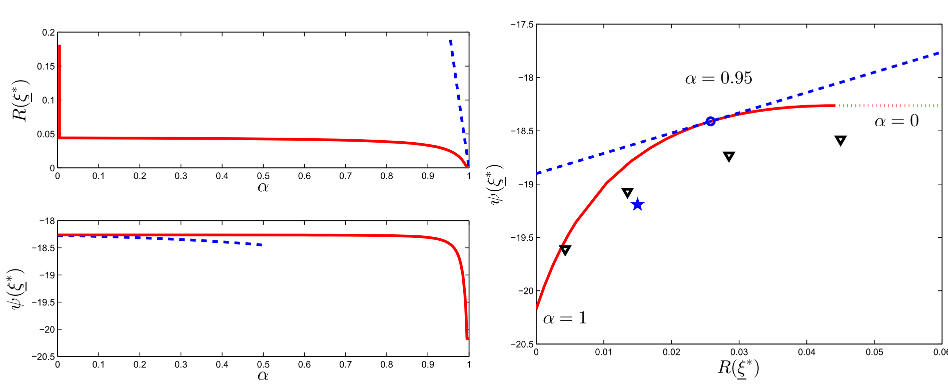

FIG. 3. *Example* 3—*Left-top*: $R(\underline{\xi}^{*(\alpha)})$ (*solid line*) *and upper bound* $4(1-\alpha)/\alpha$ (*dashed line*); *left-bottom*: $\psi(\underline{\xi}^{*(\alpha)})$ (*solid line*) *and lower bound* (2.16) (*dashed line*); *right*: $\psi(\underline{\xi}^{*(\alpha)})$ *as a function 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* (2.17) *with* $a = 1$ *and* $b = 3, 4, 6, 10$ *from right to left*.

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 $\mathscr{X}_1{}^{(0)} \simeq (0.3621, 0.5) \cup (0.6379, 1]$ and $\xi_2^{*(0)} = \mu$ on $\mathscr{X}_2{}^{(0)} \simeq [0, 0.3621) \cup (0.5, 0.6379)$. 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 $\alpha \in (0,1]$. Nondominated 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 $R(\underline{\xi}) \le R(\underline{\xi}^{*(0.95)})$.

The construction of optimal designs $\underline{\xi}^{*(\alpha)}$, together with plots similar to the one in Figure 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., [21], Chapter 9), with $[\mathrm{d}\xi_1/\mathrm{d}\mu](x)$ 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 [23], which allocates the $(n+1)$st subject to treatment 1 with probability

$$\pi_1(X_{n+1}) = \frac{d_1^a(\underline{\xi}_n, X_{n+1})}{[1-\eta_1(X_{n+1})]^b}\left(\sum_{k=1}^{2}\frac{d_k^a(\underline{\xi}_n, X_{n+1})}{[1-\eta_k(X_{n+1})]^b}\right)^{-1}, \tag{2.17}$$

with a and b some positive constants, $d_k(\underline{\xi}_n, x) = \mathrm{tr}[\mathbf{M}^{-1}(\underline{\xi}_n)\mathfrak{M}_k(x)]$, 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 [23] 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 Figure 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 the 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 α will be presented in Section 4.

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

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

$$\underline{w}(\underline{\xi}) = K \min_{k=1,\dots,K} \xi_k(\mathscr{X}). \tag{2.18}$$

The set $\Xi_\beta(\mu)$ of designs $\xi \in \Xi(\mu)$ such that $\underline{w}(\underline{\xi}) \geq \beta$ is convex, and $\underline{\xi}^{w*(\alpha,\beta)} \in \Xi_\beta(\mu)$ maximizes $H^{(\alpha)}(\cdot)$ if and only if $F_{H^{(\alpha)}}(\underline{\xi}^{w*(\alpha,\beta)}; \nu) \leq 0$ 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 $\underline{\xi}^{w*(\alpha,\beta)} = \underline{\xi}^* = (\xi_1^*, \xi_2^*) \in \Xi_\beta(\mu)$ for some $\alpha, \beta \in [0, 1]$. The following statements are equivalent*:

(i) *$\underline{\xi}^*$ is optimal, that is, it maximizes $H^{(\alpha)}(\underline{\xi})$ with respect to $\underline{\xi} \in \Xi_\beta(\mu)$;*

(ii) *there exists a constant $c = c(\alpha, \beta)$ such that $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) \geq c$ ξ_1^*-a.e. and $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) \leq c$ ξ_2^*-a.e.;*

(iii) *there exist two subsets $\mathscr{X}_1 = \mathscr{X}_1^{(\alpha,\beta)}$ and $\mathscr{X}_2 = \mathscr{X}_2^{(\alpha,\beta)}$ of $\mathscr{X}$ and a constant $c = c(\alpha, \beta)$ such that*:

 (a) *$\xi_1^* = \mu$ on $\mathscr{X}_1$ and $\xi_2^* = \mu$ on $\mathscr{X}_2$,*

 (b) *$\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) = c$ on $\mathscr{X} \setminus (\mathscr{X}_1 \cup \mathscr{X}_2)$,*

 (c) *$\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) > c$ for all $x \in \mathscr{X}_1$ and $\Delta_{12}^{(\alpha)}(\underline{\xi}^*, x) < c$ for all $x \in \mathscr{X}_2$.*

The proof is provided in the supplemental material [26]. Note that when μ 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 $\underline{\xi}^{w*(\alpha,\beta)}$.

The bounds on regret and information obtained in Section 2.4 remain valid provided that we consider designs $\underline{\xi} \in \Xi_\beta(\mu)$. In particular, using $\underline{\xi} = \underline{\xi}_\mu$ obtained for balanced random allocation (so that $\underline{\xi}_\mu \in \Xi_\beta(\mu)$ for all $\beta \in [0, 1]$), we obtain $R(\underline{\xi}^{w*(\alpha,\beta)}) \leq R(\underline{\xi}_\mu) + (1-\alpha)[\psi(\underline{\xi}^{*(0)}) - \psi(\underline{\xi}_\mu)]/\alpha$ for $\alpha > 0$, and $\psi(\underline{\xi}^{w*(\alpha,\beta)}) \geq \psi(\underline{\xi}_\mu) - \alpha R(\underline{\xi}_\mu)/(1-\alpha)$ for $\alpha < 1$.

3. Allocation with randomization.

Selection bias occurs if the experimenter is able to correctly guess next allocation in a sequential trial; see, for example, [32], Chapter 6. The bias factor B_n for n allocations is

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

(We consider the sensible guessing strategy that votes for the treatment with highest allocation probability.)

When condition (2.12) is satisfied, the allocations of Section 4.1 based on an optimal design characterized by Theorem 2.1 are deterministic (all patients with covariates in $\mathscr{X}_j$ are assigned to treatment j), and the bias factor B_n equals 1 for all n. This remains true for a response-adaptive implementation, where the sets $\mathscr{X}_j$ depend on the current estimated parameters $\hat{\theta}^n$; see Section 5: an experimenter who knows $\hat{\theta}^n$ can still predict the next allocation with certainty. Suppose now that, for each subject, with probability β we use random balanced allocation and with probability $1-\beta$ we use a predictable rule. Then, only the fraction $1-\beta$ of allocations can be guessed correctly with certainty, and $B_n \overset{\text{a.s.}}{\to} 1-\beta$ as $n \to \infty$. This section presents an extension of the results of Section 2 to this randomized framework.

3.1. *Optimal design and equivalence theorem.* For any $\underline{\xi} \in \Xi(\mu)$, we define the *uniform randomization factor* of $\underline{\xi}$ as

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

with $\mathrm{d}\xi_k/\mathrm{d}\mu$ the Radon–Nikodým derivative of ξ_k with respect to μ. [Note that $r(\underline{\xi}) \le \underline{w}(\xi)$ defined by (2.18), with equality if and only if $\xi_k = \mu/K$ for some k.] Then, if the $\xi_{n,j}$ defined by (2.3) tend to ξ_j as n tends to infinity (weak convergence), B_n satisfies $\limsup_{n\to\infty} B_n \le 1 - r(\underline{\xi})$.

Consider the maximization of $H^{(\alpha)}(\underline{\xi})$ with respect to $\underline{\xi} \in \Xi(\mu)$ under the constraint $r(\underline{\xi}) \ge \beta$, for some given $\alpha, \beta \in [0,1]$, and denote by $\underline{\xi}^{r*(\alpha,\beta)}$ an optimal design. Any admissible $\underline{\xi}$ is such that each of its components ξ_k can be decomposed as $\xi_k = (\beta/K)\mu + \tilde{\xi}_k$, where $\underline{\tilde{\xi}} = (\tilde{\xi}_1, \dots, \tilde{\xi}_K)$ belongs to $\Xi[(1-\beta)\mu]$; see (2.7). Therefore, the optimal design problem consists now in maximizing $H^{(\alpha,\beta)}(\underline{\tilde{\xi}}) = H^{(\alpha)}(\underline{\tilde{\xi}} + \beta\underline{\xi}_\mu)$ with respect to $\underline{\tilde{\xi}} \in \Xi[(1-\beta)\mu]$, with $\underline{\xi}_\mu = (\mu/K, \dots, \mu/K) \in \Xi(\mu)$. As in Section 2.2, this is a concave optimization problem over a convex set. Optimal designs $\underline{\xi}^{r*(\alpha,\beta)} = \underline{\tilde{\xi}}^* + \beta\underline{\xi}_\mu$ are still characterized by Theorem 2.1, with the following slight modifications: the statement in (ii) is now valid $\tilde{\xi}_i^*$-a.e.; $\sum_{i\in\mathscr{J}_t} \tilde{\xi}_i^* = (1-\beta)\mu$ 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 $\mathscr{X}_1$ and $\xi_2^* = (1-\beta/2)\mu$ on $\mathscr{X}_2$. Note that $H^{(\alpha,\beta)}(\cdot)$ is differentiable at any $\underline{\tilde{\xi}} \in \Xi[(1-\beta)\mu]$ when $\beta > 0$, since we have assumed that $\mathbf{M}(\underline{\xi}_\mu) \in \mathbb{M}^+$.

REMARK 3.1. The case of unbalanced randomization, with different values of β for different k, could be treated in the same way, with now $\xi_k = \beta_k\mu + \tilde{\xi}_k$, where $\beta_k \ge 0$ for all k, $\sum_{k=1}^K \beta_k = B \in [0,1]$ and $\underline{\tilde{\xi}} = (\tilde{\xi}_1, \dots, \tilde{\xi}_K) \in \Xi[(1-B)\mu]$. Also, similar developments can be made when the randomization is not uniform over $\mathscr{X}$ and the constraint on $\underline{\xi}$ takes the form $(\mathrm{d}\xi_k/\mathrm{d}\mu)(x) \ge \beta(x)/K$ for $k = 1, \dots, K$, with $\beta(x) \in [0, 1]$ for all $x \in \mathscr{X}$. In that case, each ξ_k can be decomposed into $\xi_k = \nu_k + \tilde{\xi}_k$, where $(\mathrm{d}\nu_k/\mathrm{d}\mu)(x) = \beta(x)/K$ for all k and $\underline{\tilde{\xi}} = (\tilde{\xi}_1, \dots, \tilde{\xi}_K) \in \Xi(\tilde{\mu}_\beta)$ with $(\mathrm{d}\tilde{\mu}_\beta/\mathrm{d}\mu)(x) = 1 - \beta(x)$.

3.2. *Bounds on optimal regret and information.* When $\beta > 0$, due to the linearity in $\underline{\xi}$ of $R(\underline{\xi})$, the regret for an optimal design $\underline{\xi}^{r*(\alpha,\beta)} = \underline{\tilde{\xi}}^* + \beta\underline{\xi}_\mu$ can be decomposed into two parts, $R(\underline{\xi}^{r*(\alpha,\beta)}) = R(\underline{\tilde{\xi}}^*) + \beta R(\underline{\xi}_\mu)$, where $R(\underline{\tilde{\xi}}^*)$ satisfies an inequality similar to (2.13). Therefore, for $\alpha > 0$, $R(\underline{\xi}^{r*(\alpha,\beta)}) \le [(1 - \alpha)/\alpha]\,\mathrm{tr}[\nabla\psi(\underline{\xi}^{r*(\alpha,\beta)})\mathbf{M}(\underline{\tilde{\xi}}^*)] + \beta R(\underline{\xi}_\mu)$. When $\Psi(\mathbf{M}) = \log\det(\mathbf{M})$, we obtain

$$R(\underline{\xi}^{r*(\alpha,\beta)}) \le p\frac{1-\alpha}{\alpha} + \beta R(\underline{\xi}_\mu) - \beta\frac{1-\alpha}{\alpha}\,\mathrm{tr}[\mathbf{M}^{-1}(\underline{\xi}^{r*(\alpha,\beta)})\mathbf{M}(\underline{\xi}_\mu)].$$

Using the rough bound $\mathrm{tr}[\mathbf{M}^{-1}(\underline{\xi}^{r*(\alpha,\beta)})\mathbf{M}(\underline{\xi}_\mu)] > \mathrm{tr}\{[K\mathbf{M}(\underline{\xi}_\mu)]^{-1}\mathbf{M}(\underline{\xi}_\mu)\} = p/K$, we get

$$R(\underline{\xi}^{r*(\alpha,\beta)}) < p\left(1 - \frac{\beta}{K}\right)\frac{1-\alpha}{\alpha} + \beta R(\underline{\xi}_\mu).$$

We can also obtain an inequality similar to (2.15), provided that we consider $\underline{\xi} \in \Xi(\mu)$ such that $r(\underline{\xi}) \ge \beta$.

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 $r(\underline{\xi}) \ge \beta$. As in Section 2.4, we have $\psi(\underline{\xi}^{r*(\alpha,\beta)}) \ge \psi(\underline{\xi}) + \alpha[\phi(\underline{\xi}) - \phi(\underline{\xi}^{r*(\alpha,\beta)})]/(1-\alpha)$ and, therefore,

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

where we have used the linearity of $\phi(\underline{\xi})$ with respect to $\underline{\xi}$. Using (2.11), we obtain in particular

$$\psi(\underline{\xi}^{r*(\alpha,\beta)}) \ge \psi(\underline{\xi}_\mu) - (1-\beta)\frac{\alpha}{1-\alpha}R(\underline{\xi}_\mu),$$

which can be compared with (2.16).

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

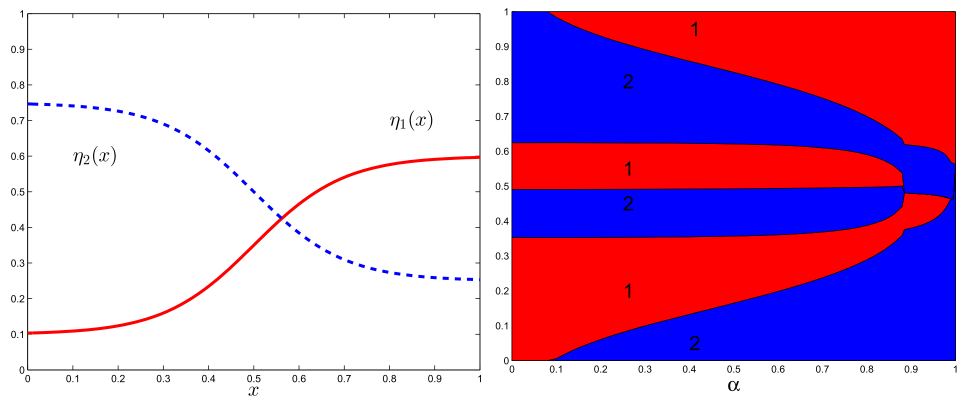

FIG. 4. *Example* 2—*Left*: $\eta_1(x)$ (*solid line*) *and* $\eta_2(x)$ (*dashed line*) *as functions of* x; *right*: *subsets* $\mathscr{X}_1^{(\alpha)}$ *and* $\mathscr{X}_2^{(\alpha)}$ *of Corollary* 2.1(ii) *for* $\alpha \in [0, 1]$ *and* $\beta = 0.2$.

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 nth 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

$$\pi_k(X_{n+1}) = \mathrm{Prob}(T_{n+1,k} = 1 | X_{n+1}, \mathscr{F}_n), \qquad k = 1, \dots, K, \tag{4.1}$$

with $\mathscr{F}_n$ the filtration $\sigma(\mathbf{T}_1, \dots, \mathbf{T}_n, X_1, \dots, X_n)$. Denote $\mathbf{N}_n = \sum_{i=1}^n \mathbf{T}_i$, so that its kth 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 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. If $\underline{\xi}^*$ is known, a straightforward construction ensuring $\xi_{n,k} \stackrel{\text{a.s.}}{\to} \xi_k^*$ (weak convergence) consists in sampling according to ξ_k^*; see (4.2). 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. In that case, we 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

$$\pi_k^*(X_n) = \frac{\mathrm{d}\xi_k^*}{\mathrm{d}\mu}(X_n), \qquad n \geq 1, \tag{4.2}$$

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) corresponds to the following generalized biased-coin design:

$$\pi_k^*(X_n) = \begin{cases} 1 - \dfrac{(K-1)\beta}{K} & \text{if } G_k^{(\alpha)}(\underline{\xi}^*, X_n) = \max\limits_{j=1,\dots,K} G_j^{(\alpha)}(\underline{\xi}^*, X_n), \\ \dfrac{\beta}{K} & \text{otherwise} \end{cases} \tag{4.3}$$

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

$$\mathbf{N}/n \overset{\text{a.s.}}{\to} \boldsymbol{\rho}^* = \boldsymbol{\rho}(\underline{\xi}^*) = (\xi_1^*(\mathscr{X}), \dots, \xi_K^*(\mathscr{X}))^\top \tag{4.4}$$

from the strong law of large numbers (SLLN), and

$$\sqrt{n}(\mathbf{N}/n - \boldsymbol{\rho}^*) \overset{\text{d}}{\to} \mathscr{N}(\mathbf{0}, \Sigma^*), \qquad \text{with } \Sigma^* = \text{diag}(\boldsymbol{\rho}^*) - \boldsymbol{\rho}^* \boldsymbol{\rho}^{*\top}, \tag{4.5}$$

from the central limit theorem (CLT). The proportions N_n/n also satisfy the law of the iterated logarithm. 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 $R_n \overset{\text{a.s.}}{\to} R(\underline{\xi}^*)$ and $\mathbf{M}_n \overset{\text{a.s.}}{\to} \mathbf{M}(\underline{\xi}^*)$, where $\underline{\xi}_n^* = (\xi_{n,1}^*, \dots, \xi_{n,K}^*)$ with $\xi_{n,k}^*$ the empirical measure (2.3). We thus have $\psi(\underline{\xi}_n^*) \overset{\text{a.s.}}{\to} \psi(\underline{\xi}^*)$ and $H^{(\alpha)}(\underline{\xi}_n^*) \overset{\text{a.s.}}{\to} H^{(\alpha)}(\underline{\xi}^*)$. The values of $R(\underline{\xi}_n^*)$ and $\psi(\underline{\xi}_n^*)$ also obey the CLT; direct calculations show that $\sqrt{n}[R(\underline{\xi}_n^*) - R(\underline{\xi}^*)] \overset{\text{d}}{\to} \mathscr{N}(0, V_R^*)$ and (using the delta method) that $\sqrt{n}[\psi(\underline{\xi}_n^*) - \psi(\underline{\xi}^*)] \overset{\text{d}}{\to} \mathscr{N}(0, V_\psi^*)$, with

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

$$V_\psi^* = \sum_{k=1}^K \mathsf{P}_{\xi_k^*} \operatorname{tr}^2[\nabla\psi(\underline{\xi}^*)\mathfrak{M}_k(\cdot)] - \operatorname{tr}^2[\nabla\psi(\underline{\xi}^*)\mathbf{M}(\underline{\xi}^*)].$$

Although attractive from a theoretical viewpoint, (4.2) has the inconvenient that it requires the knowledge of μ and relies on the prior construction of an optimal design $\underline{\xi}^*$. Note that this construction may be difficult when $\mathscr{X}$ and μ satisfy H1b with $d = \dim(X) > 1$. Moreover, 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 θ (see Section 5), which means that an oracle providing $\underline{\xi}^*(\theta)$ for any θ should be available. In the next section, we consider an allocation rule $\hat{\pi}_k(X_n)$ that asymptotically samples from $\underline{\xi}^*$ without requiring neither the explicit construction of $\underline{\xi}^*$ nor the knowledge of μ.

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, that is, adaptive also with respect to previous allocations, and uses allocation probabilities given by (for n larger than some $n_0 > 0$)

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

where $\ell_n(x) = |\{j \in \{1,\dots,K\} : G_j^{(\alpha)}(\hat{\underline{\xi}}_n, x) = \max_{k=1,\dots,K} G_k^{(\alpha)}(\hat{\underline{\xi}}_n, x)\}|$. 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 below indicates that when $\beta > 0$, $\hat{\underline{\xi}}_n$ generated by (4.6) has the same asymptotic information and regret values as $\underline{\xi}^{r*(\alpha,\beta)}$ of Section 3.

THEOREM 4.1. *Under* H1a *or* (H1b, H2, H2′), *for any* $\beta \in (0, 1]$ *and* $\alpha \in [0, 1]$, *the allocation rule* (4.6) *satisfies*

$$
H^{(\alpha)}(\hat{\underline{\xi}}_n) \overset{\text{a.s.}}{\to} H^{(\alpha)}\big(\underline{\xi}^{r*(\alpha,\beta)}\big), \qquad n \to \infty,
$$

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* $r(\underline{\xi}) \ge \beta$; *see* (3.1). *Moreover*, $\mathbf{M}(\hat{\underline{\xi}}_n) \overset{\text{a.s.}}{\to} \mathbf{M}(\underline{\xi}^{r*(\alpha,\beta)})$, $\psi(\hat{\underline{\xi}}_n) \overset{\text{a.s.}}{\to} \psi(\underline{\xi}^{r*(\alpha,\beta)})$, *and also* $R(\hat{\underline{\xi}}_n) \overset{\text{a.s.}}{\to} R(\underline{\xi}^{r*(\alpha,\beta)})$ *if* $\alpha > 0$.

When $\alpha = 0$ and $\Psi(\mathbf{M}) = \log\det(\mathbf{M})$, (4.6) corresponds to the sequential construction of a D-optimal design in $\Xi(\mu)$; see [1–3]. Notice 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 (4.6) does not enter the general framework considered in [8]. The proof of Theorem 4.1 is presented in the supplemental material [26]. The assumption that $\mathscr{X}$ 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 [28], Theorem 9.

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 [36]: either $H^{(\alpha)}(\hat{\underline{\xi}}_n) \to H^{(\alpha)}(\underline{\xi}^{r*(\alpha,\beta)})$, or $\liminf_{n\to\infty} \Psi[\mathbf{M}(\hat{\underline{\xi}}_n)] = -\infty$ when $\Psi(\cdot)$ is one of the criteria (2.6) ($\liminf_{n\to\infty} \Psi[\mathbf{M}(\hat{\underline{\xi}}_n)] = 0$ for their positively homogeneous versions). However, since the X_i are i.i.d. in $\mathscr{X}$ and $\mathbf{M}(\underline{\xi}_\mu) \in \mathbb{M}^+$, one can force the second event to have zero probability. For $\alpha \in [0, 1)$ and $\beta = 0$, we modify the rule (4.6) through the introduction of a lower bound $\psi_*^{(\alpha)}$ on $\psi(\underline{\xi}^{*(\alpha)})$, obtained for instance from (2.16). For each $n \geq n_0$, we allocate the $(n+1)$st subject to treatment k with probability $1/K$ if $\psi(\hat{\underline{\xi}}_n) < \min\{\psi_*^{(\alpha)}, \psi(\underline{\xi}_\mu)\}$ 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 (4.6). Then, for $\alpha < 1$ and $\beta = 0$, the second-order derivative of $\psi(\cdot)$ at $\hat{\underline{\xi}}_n$ is bounded from below (a.s.), and the empirical design $\hat{\underline{\xi}}_n$ 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 (4.6) converge (a.s.) to those obtained with the rule (4.2) based on an oracle optimal design. Under the conditions mentioned at the end of 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 Σ^* obtained with $\pi_k^*(X_{n+1})$ [see (4.5)], a phenomenon that resembles the improved treatment balance obtained by the method of [27] generalizing [16]; see also [4]. Assuming that θ is known, the doubly-adaptive designs of [38], which extend the approach of [22] to the presence of covariates, are able to yield (in the limit) a reduction of Σ^* to

$$\mathrm{Var}(\underline{\pi}^*) = \mathsf{P}_\mu\big[\big(\underline{\pi}^*(\cdot) - \boldsymbol{\rho}^*\big)\big(\underline{\pi}^*(\cdot) - \boldsymbol{\rho}^*\big)^\top\big],$$

where $\underline{\pi}^*(x) = (\pi_1^*(x), \ldots, \pi_K^*(x))^\top$ with $\pi_k^*(x)$ and $\boldsymbol{\rho}^*$ respectively given by (4.2) and (4.4). Note that $\mathrm{Var}(\underline{\pi}^*)/n$ corresponds to the asymptotic covariance matrix of the ML estimator $\hat{\boldsymbol{\rho}}^n = (1/n)\sum_{i=1}^n \underline{\pi}^*(X_i)$—under the assumption that the $\underline{\pi}^*(X_i)$ are i.i.d. When $\beta = 0$ and (2.12) is satisfied, so that $\pi_k^*(x) \in \{0, 1\}$ for all x, then $\mathrm{Var}(\underline{\pi}^*) = \Sigma^*$. Numerical simulations show that the rule $\hat{\pi}_k(X_{n+1})$ given by (4.6), which is design-adaptive, achieves a smaller asymptotic variance than $\mathrm{Var}(\underline{\pi}^*)$ for the proportions $\mathbf{N}/n$. This is illustrated in Section 4.3.1 which continues Example 3.

REMARK 4.1. (i) The approach used in [38] 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 $\mathbf{M}(\hat{\underline{\xi}}_n)$ and not only the proportions $\mathbf{N}/n$.

(ii) The fact that (4.6) yields a smaller asymptotic variance than $\mathrm{Var}(\underline{\pi}^*)$ for the proportions $\mathbf{N}/n$ indicates than the efficient designs in Definition 2.1 of [38] are not of asymptotic minimum variability within the class of covariate-adjusted designs satisfying condition (2.2) in the same paper. On the other hand, the notion of efficiency in [38], Definition 2.1, remains valid when restricted to the class of CARA rules of the form $\underline{\pi}(X_{n+1}, \hat{\theta}^n)$.

REMARK 4.2. When $\underline{\xi}^*$ is not unique, one may wonder what is the limiting design for (4.6). Numerical simulations with Example 1 indicate convergence to a unique design, whatever the initialization of the sequential procedure (with n_0 arbitrarily large). Further developments are required to investigate if some stability properties of (4.6) around an optimal design $\underline{\xi}^*$ would permit to characterize which particular optimal designs can be reached in the limit.

4.3. *Examples*.

4.3.1. *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 1000 simulations of allocation rules (4.3) and (4.6) with $n = 5000$ subjects (we use $n_0 = 4$, with two initial assignments of each treatment). Empirical distributions are smoothed with a normal kernel density estimator, using Silverman's rule for bandwidth selection.

Figure 5 shows the empirical distributions of $\sqrt{n}(N_{n,k}/n - \boldsymbol{\rho}_k^*)$, for $k = 1$ (top) and $k = 2$ (bottom). The dashed-line curves are for (4.3) and show good agreement with the asymptotic distributions $\mathscr{N}(0, {\Sigma_{kk}^*}^2)$ (solid-line); the dotted-line curves are for (4.6) which exhibits smaller variability around the optimal proportions $\boldsymbol{\rho}^*$.

4.3.2. *Example* 4 (*continued*). Consider again the situation of Example 4, (see Section 3.3) with $\alpha = 0.7$ and $\beta = 0.2$. Figure 6(left) presents histograms of $\hat{\xi}_{n,1}$ and $\hat{\xi}_{n,2}$ obtained with the allocation rule (4.6) (with $n = 2 \times 10^4$ and $n_0 = 4$—two initial assignments of each treatment). Note the good agreement with the optimal design $\underline{\xi}^{r*(\alpha,\beta)}$ presented in Figure 4(right), where $\xi_1^* = (1 - \beta/2)\mu$ on $\mathscr{X}_1 \simeq (0.237, 0.368) \cup (0.495, 0.610) \cup (0.7525, 1]$.

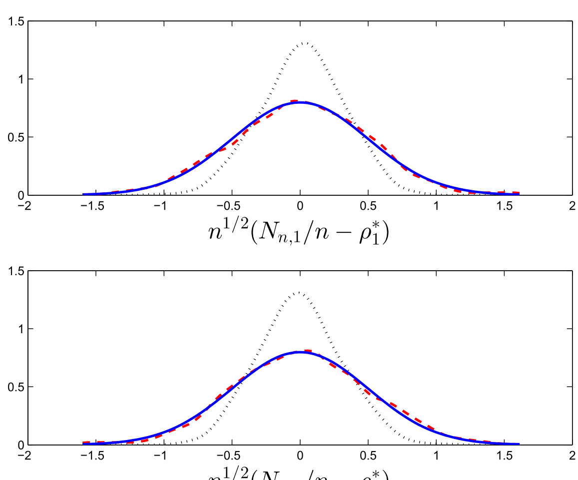

FIG. 5. *Example* 3 ($\alpha = 0.7$, $\beta = 0$, $n = 5000$, 1000 *repetitions*): *empirical distributions of* $\sqrt{n}(N_{n,k}/n - \boldsymbol{\rho}_k^*)$, $k = 1, 2$, *for* (4.3) (*dashed-lines*) *and* (4.6) (*dotted lines*), *the solid-lines correspond to* $\mathscr{N}(0, {\Sigma_{kk}^*}^2)$.

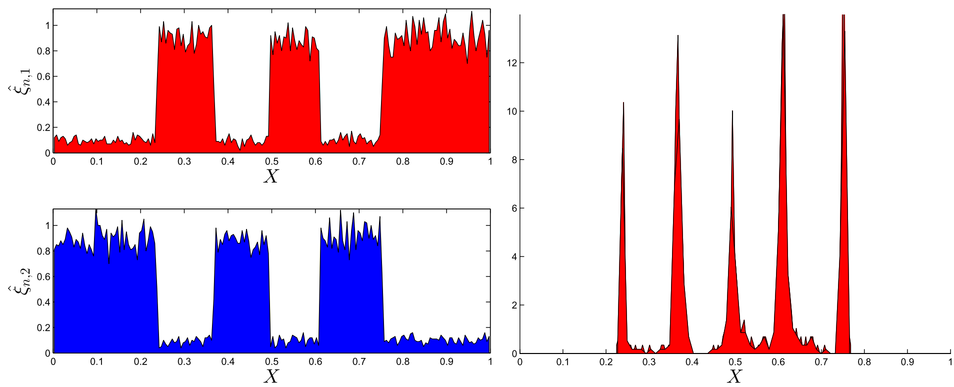

FIG. 6. *Example* 4 ($\alpha = 0.7$, $\beta = 0.2$)—*Left*: *histograms of* $\hat{\xi}_{n,1}$ *and* $\hat{\xi}_{n,2}$, $n = 2 \times 10^4$; *right*: *histogram of covariates for which* (4.3) *would disagree from* (4.6).

In fact, for most assignments the rule (4.6) agrees with (4.3) which samples from $\underline{\xi}^{r*(\alpha,\beta)}$: different allocations occur essentially for values of x near the end-points of the intervals that define $\mathscr{X}_1$; see Figure 6(right) for a histogram of the values of X where (4.3) would give a treatment different from that given by (4.6). Figure 7(left) shows the number N_n^D of disagreements with (4.3) when (4.6) is used in a sequence of length n. $G_j^{(\alpha)}(\hat{\underline{\xi}}_n, x)$ 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 7(right) presents the evolution of $R(\hat{\underline{\xi}}_n)$ (top) and $H(\hat{\underline{\xi}}_n)$ (bottom) as functions of n for the rule (4.6). Convergence to the optimal values (indicated by dashed lines) is reason-

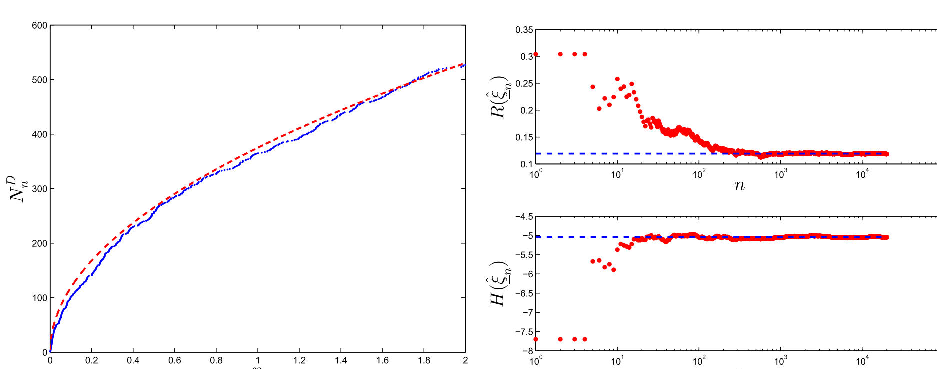

FIG. 7. *Example* 4 ($\alpha = 0.7$, $\beta = 0.2$)—*Left*: *number of disagreements with* (4.3) *when* (4.6) *is used in a sequence of length n* (*the curve in dashed line corresponds to* $3.75\sqrt{n}$); *right*: $R(\hat{\underline{\xi}}_n)$ (*top*) *and* $H(\hat{\underline{\xi}}_n)$ (*bottom*) *as functions of n for allocation rule* (4.6), *optimal values are indicated in dashed lines*.

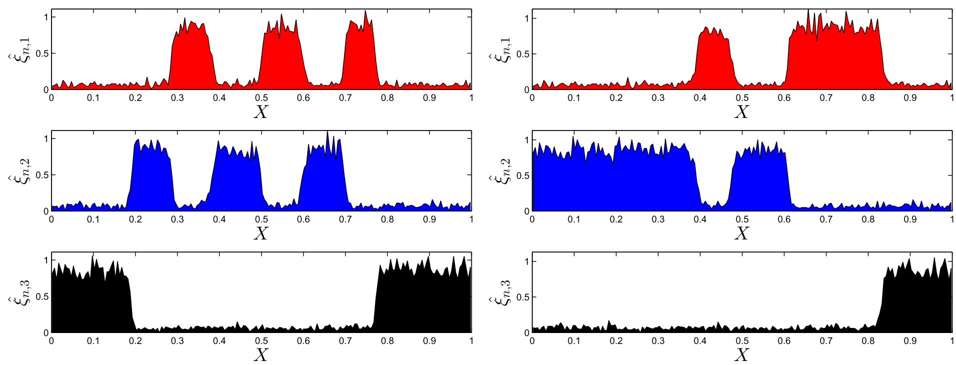

FIG. 8. *Example* 5—*Histograms of* $\hat{\xi}_{n,k}$, $k = 1, 2, 3$ ($n = 2 \times 10^4$) *for* $\alpha = 0.7$, $\beta = 0.2$ (*left*) *and* $\alpha = 0.9$, $\beta = 0.2$ (*right*).

ably fast; the figure is quasi identical when (4.3) is used with the same sequence of covariates.

4.3.3. *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 (4.6) ($n = 2 \times 10^4$ and $n_0 = 6$) are presented in Figure 8, respectively, for $\alpha = 0.7$, $\beta = 0.2$ and $\alpha = 0.9$, $\beta = 0.2$. For α or β large enough, the optimal design is such that $\xi_3^* = (\beta/3)\mu$, that is, the random component of the design gives enough precision for the estimation of θ_3 (the placebo effect), taking the poor efficacy of this treatment into account.

5. Further extensions and developments. We have proposed sequential allocation rules that target optimal compromise strategies, in the sense that the asymptotic regret and information values are nondominated, contrasting with ad hoc rules whose asymptotic regret and information can both be improved. These results can be extended in various directions.

Extension to response-adaptive rules. As usual in nonlinear situations, optimal designs $\underline{\xi}_\theta^{*(\alpha)}$ depend on the unknown value θ of the model parameters. Here, we only considered locally optimum design, where θ is set to a given nominal value θ^0. In a response-adaptive implementation, when assigning the $(n+1)$st subject, θ^0 can be replaced by $\hat{\theta}^n$, the current ML estimator of θ 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 designs 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 [21], page 155, [39], Section 4. 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 $\underline{\xi}_{\hat{\theta}^n}^{*(\alpha)}$ for each n, with $\hat{\theta}^n$ in a compact subset Θ of $\mathbb{R}^p$ containing admissible θ. When $\alpha < 1$, developments similar to those in [39] can be used to prove the strong consistency and asymptotic normality of $\hat{\theta}^n$, with $\sqrt{n}(\hat{\theta}^n - \theta) \xrightarrow{\mathrm{d}} \mathscr{N}(\mathbf{0}, \mathbf{M}^{-1}(\underline{\xi}_{\theta}^{*(\alpha)}; \theta))$ under rather standard regularity assumptions. Condition A of [39] is not satisfied since $\mu\{x : \pi_k^*(x; \theta) = 0\} > 0$ when $\pi_k^*(x; \theta)$ is given by (4.2), but it may be counterbalanced by the assumption that all optimal information matrices $\mathbf{M}(\underline{\xi}_{\theta}^{*(\alpha)}; \theta)$, $\theta \in \Theta$, are nonsingular. The asymptotic normality of allocation proportions $\mathbf{N}/n$, information and regret, can also be proved, similarly to Section 4.1 but with larger variances. A reduction of variance based on the approach in [38] might then be considered, 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 $\widehat{g}_k^{(\alpha)}(x, \theta)$ of the functions $g_k^{(\alpha)}(x, \theta) = G_k^{(\alpha)}(\underline{\xi}_{\theta}^{*(\alpha)}, x; \theta)$ defined on $\mathscr{X} \times \Theta$, based on the knowledge of $x \mapsto g_k^{(\alpha)}(x, \theta^{(\ell)})$ for a few $\theta^{(\ell)}$, and then compute the K allocation probabilities $\widehat{\pi}_k^*(X_{n+1}; \hat{\theta}^n)$ for every $\hat{\theta}^n$ by using Theorem 2.1 and the approximations $\widehat{g}_k^{(\alpha)}(x, \hat{\theta}^n)$.

Alternatively, one may consider a response-adaptive version of (4.6), with $G_k^{(\alpha)}(\underline{\xi}_n, x; \hat{\theta}^n)$ substituted for $G_k^{(\alpha)}(\underline{\xi}_n, x; \theta^0)$. One can show that the strong consistency of $\hat{\theta}^n$ and its asymptotic normality are preserved under suitable regularity assumptions, with $\sqrt{n}(\hat{\theta}^n - \theta) \xrightarrow{\mathrm{d}} \mathscr{N}(\mathbf{0}, \mathbf{M}^{-1}(\underline{\xi}_{\theta}^{*(\alpha)}; \theta))$. 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 $R_n = (1/n) \sum_{i=1}^n \mathbb{I}\{\eta_{k_i}(X_i) \neq \eta_*(X_i)\} = 1 - \sum_{k=1}^K \xi_{n,k}\{x : \eta_k(x) = \eta_*(x)\}$. This means replacing the responses η_k by the indicator functions $\eta'_k(x) = \mathbb{I}\{\eta_k(x) = \eta_*(x)\}$ in $\phi(\underline{\xi})$; see (2.9), and $G_k^{(\alpha)}(\underline{\xi}, x)$, see (2.10). Also, when the η_k correspond to success probabilities for binary responses, one may enforce individual ethics by increasing the penalty for not using the best treatment, and consider $R_n = \sum_{l=1}^n \mathsf{P}_{\xi_{n,k}}[\eta_k^q(\cdot) - \eta_*^q(\cdot)]$ with $q < 0$. The definition of regret may also be extended to account for efficacy and toxicity in the case of bivariate binary responses. Most of the results presented remain valid provided that $R(\underline{\xi})$ is linear in $\underline{\xi}$.

Other measures of information. Here, we have considered a global measure of information, accounting for the precision of the estimation of all components of θ. One might alternatively consider an information criterion related to a specific feature of the model responses, like the treatment difference, possibly in a particular range of the covariates. For instance, in Example 1, one may be interested in the estimation of $b_2 - b_1$, considering a_0 as a nuisance parameter. This would provide an extension of the approach in [1–3, 5] to the case of generalized linear models with attention to ethical cost.

Choice of α. When θ is fixed, the behaviours of $R(\underline{\xi}^{*(\alpha)})$ and $\psi(\underline{\xi}^{*(\alpha)})$ as functions of α (see Figure 3) allow an iterative choice of α to target a prescribed risk or information value. Also, the bounds given in Sections 2.4 and 3.2 may guide the selection of a suitable α.

In the sequential framework considered in Section 4.2, an alternative solution would be to specify a target τ on the regret, and maximize information under the constraint that $R(\underline{\xi}) \le \tau$, with associated Lagrangian $\mathscr{L}(\underline{\xi}, C) = \psi(\underline{\xi}; \theta) - C[R(\underline{\xi}; \theta) - \tau]$; see Section 2.1. In a CARA scheme, one might then let the Lagrange coefficient C vary with n as $C_{n+1} = \max\{0, C_n + \gamma[R(\hat{\underline{\xi}}_n; \hat{\theta}^n) - \tau]\}$, with γ some positive constant. This is equivalent to letting α 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 $n \to \infty$ 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, $\mathsf{E}\{R_n\} = O(\log(n)/n)$; see, for instance, [11, 20, 25]. Also, is is tempting to relate α_n to the precision of the estimation of θ in order to focus on allocation to the best treatment when there is enough evidence of which treatment is best, with some similarities with the approach in [35], Section 6, that gradually shifts emphasis from model discrimination to parameter estimation.

SUPPLEMENTARY MATERIAL

Supplement to "Information-regret compromise in covariate-adaptive treatment allocation" (DOI: 10.1214/16-AOS1518SUPP; .pdf). In this Supplement, we give the proofs of Theorems 2.1, 2.2 and 4.1.

REFERENCES

[1] ATKINSON, A. C. (1982). Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika* **69** 61–67. MR0655670

[2] ATKINSON, A. C. (1999). Optimum biased-coin designs for sequential treatment allocation with covariate information. *Stat. Med.* **18** 1741–1752.

[3] ATKINSON, A. C. (2002). The comparison of designs for sequential clinical trials with covariate information. *J. Roy. Statist. Soc. Ser. A* **165** 349–373. MR1904822

[4] ATKINSON, A. C. (2014). Selecting a biased-coin design. *Statist. Sci.* **29** 144–163. MR3201860

[5] ATKINSON, A. C. and BISWAS, A. (2005). Bayesian adaptive biased-coin designs for clinical trials with normal responses. *Biometrics* **61** 118–125. MR2135851

[6] BALDI ANTOGNINI, A. and GIOVAGNOLI, A. (2010). Compound optimal allocation for individual and collective ethics in binary clinical trials. *Biometrika* **97** 935–946. MR2746162

[7] BALDI ANTOGNINI, A. and GIOVAGNOLI, A. (2015). *Adaptive Designs for Sequential Treatment Allocation*. CRC Press, Boca Raton.

[8] BALDI ANTOGNINI, A. and ZAGORAIOU, M. (2015). On the almost sure convergence of adaptive allocation procedures. *Bernoulli* **21** 881–908. MR3338650

[9] BALL, F. G., SMITH, A. F. M. and VERDINELLI, I. (1993). Biased coin designs with a Bayesian bias. *J. Statist. Plann. Inference* **34** 403–421. MR1210443

[10] BISWAS, A., PARK, E. and BHATTACHARYA, R. (2012). Covariate-adjusted response-adaptive designs for longitudinal treatment responses: PEMF trial revisited. *Stat. Methods Med. Res.* **21** 379–392. MR3190616

[11] CAPPÉ, O., GARIVIER, A., MAILLARD, O.-A., MUNOS, R. and STOLTZ, G. (2013). Kullback–Leibler upper confidence bounds for optimal sequential allocation. *Ann. Statist.* **41** 1516–1541.

[12] COOK, R. D. and WONG, W. K. (1994). On the equivalence of constrained and compound optimal designs. *J. Amer. Statist. Assoc.* **89** 687–692. MR1294092

[13] DRAGALIN, V. and FEDOROV, V. (2006). Adaptive designs for dose-finding based on efficacy-toxicity response. *J. Statist. Plann. Inference* **136** 1800–1823. MR2255597

[14] DRAGALIN, V., FEDOROV, V. and WU, Y. (2008). Adaptive designs for selecting drug combinations based on efficacy-toxicity response. *J. Statist. Plann. Inference* **138** 352–373.

[15] DRAGALIN, V., FEDOROV, V. and WU, Y. (2008). Two-stage design for dose-finding that accounts for both efficacy and safety. *Stat. Med.* **27** 5156–5176.

[16] EFRON, B. (1971). Forcing a sequential experiment to be balanced. *Biometrika* **58** 403–417. MR0312660

[17] FEDOROV, V. V. (1989). Optimal design with bounded density: Optimization algorithms of the exchange type. *J. Statist. Plann. Inference* **22** 1–13. MR0996795

[18] FEDOROV, V. V. and LEONOV, S. L. (2014). *Optimal Design for Nonlinear Response Models*. CRC Press, Boca Raton, FL.

[19] GOLDENSHLUGER, A. and ZEEVI, A. (2011). A note on performance limitations in bandit problems with side information. *IEEE Trans. Inform. Theory* **57** 1707–1713.

[20] GOLDENSHLUGER, A. and ZEEVI, A. (2013). A linear response bandit problem. *Stoch. Syst.* **3** 230–261. MR3353472

[21] HU, F. and ROSENBERGER, W. F. (2006). *The Theory of Response-Adaptive Randomization in Clinical Trials*. Wiley, Hoboken, NJ. MR2245329

[22] HU, F. and ZHANG, L.-X. (2004). Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. *Ann. Statist.* **32** 268–301. MR2051008

[23] HU, J., ZHU, H. and HU, F. (2015). A unified family of covariate-adjusted response-adaptive designs based on efficiency and ethics. *J. Amer. Statist. Assoc.* **110** 357–367. MR3338508

[24] HU, Y. and HU, F. (2012). Asymptotic properties of covariate-adaptive randomization. *Ann. Statist.* **40** 1794–1815. MR3015044

[25] LAI, T. L. and ROBBINS, H. (1985). Asymptotically efficient adaptive allocation rules. *Adv. in Appl. Math.* **6** 4–22.

[26] METELKINA, A. and PRONZATO, L. (2017). Supplement to "Information-regret compromise in covariate-adaptive treatment allocation." DOI:10.1214/16-AOS1518SUPP.

[27] POCOCK, S. J. and SIMON, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* **31** 103–115.

[28] PRONZATO, L. (2006). On the sequential construction of optimum bounded designs. *J. Statist. Plann. Inference* **136** 2783–2804.

[29] PRONZATO, L. (2010). Penalized optimal designs for dose-finding. *J. Statist. Plann. Inference* **140** 283–296. MR2568139

[30] PUKELSHEIM, F. (1993). *Optimal Experimental Design*. Wiley, New York.

[31] RABIE, H. and FLOURNOY, N. (2013). Optimal designs for contingent response models with application to toxicity-efficacy studies. *J. Statist. Plann. Inference* **143** 1371–1379. MR3055754

[32] ROSENBERGER, W. F. and LACHIN, J. M. (2002). *Randomization in Clinical Trials. Theory and Practice*. Wiley, New York. MR1914364

[33] ROSENBERGER, W. F., SVERDLOV, O. and HU, F. (2012). Adaptive randomization for clinical trials. *J. Biopharm. Statist.* **22** 719–736. MR2931067

[34] SAHM, M. and SCHWABE, R. (2001). A note on optimal bounded designs. In *Optimum Design* 2000 (*Cardiff*). *Nonconvex Optim. Appl.* **51** 131–140. Kluwer Academic, Dordrecht. MR1829425

[35] WIENS, D. P. (2009). Robust discrimination designs. *J. R. Stat. Soc., B* **71** 805–829.

[36] WU, C. F. J. and WYNN, H. P. (1978). The convergence of general step–length algorithms for regular optimum design criteria. *Ann. Statist.* **6** 1273–1285. MR0523762

[37] WYNN, H. P. (1982). Optimum submeasures with applications to finite population sampling. In *Statistical Decision Theory and Related Topics III* (S. S. Gupta and J. O. Berger, eds.). *Proc.* 3*rd Purdue Symp.* **2** 485–495. Academic Press, New York.

[38] ZHANG, L.-X. and HU, F. (2009). A new family of covariate-adjusted response-adaptive designs and their properties. *Appl. Math. J. Chinese Univ.* **24** 1–13.

[39] ZHANG, L.-X., HU, F., CHEUNG, S. H. and CHAN, W. S. (2007). Asymptotic properties of covariate-adjusted response-adaptive designs. *Ann. Statist.* **35** 1166–1182. MR2341702

LABORATOIRE I3S, UNS-CNRS
BÂT. EUCLIDE, LES ALGORITHMES, BP 121
2000 ROUTE DES LUCIOLES
06903 SOPHIA ANTIPOLIS CEDEX
FRANCE
E-MAIL: metelkin@unice.fr
Luc.Pronzato@cnrs.fr