Statistica Sinica: Supplement

A STATISTICAL APPROACH TO ADAPTIVE PARAMETER TUNING IN NATURE-INSPIRED OPTIMIZATION AND OPTIMAL SEQUENTIAL DESIGN OF DOSE-FINDING TRIALS

Kwok Pui Choi, Tze Leung Lai, Xin T. Tong and Weng Kee Wong

National University of Singapore, Stanford University and UCLA

Supplementary Material

This is the supplementary material for the article "A statistical approach to adaptive parameter tuning in nature-inspired optimization and optimal sequential design of dose-finding trials". All sections and equations refer to the main text.

S1 Proof of Theorem 2

Section 1 of Chan and Lai (2006) reviews the literature on selection and ranking, and in particular the δ-difference zone approach (Bechhofer, Kiefer, and Sobel, 1968), which we now apply to multi-armed bandits. For the one-parameter exponential family, UCB of an individual arm based on observations up to time t from that arm is derived by inverting a GLR test; see Lai (1987, Section 2), where $\log \alpha^{-1} \sim \log B$ for the confidence level $1-\alpha$ under δ-indifference and B is the sample size in Theorem 1. The asymptotic lower bound for the regret that we have reviewed in Section 2.3 can be translated in terms of the total sampling cost C_T since $\inf_{\mathbf{x}} g(\mathbf{x}) > 0$ is assumed in Theorem 2. Extension of this argument to the multiparameter exponential

family, which is the setting of Theorems 1 and 2, is straightforward, particularly since we use ϵ-greedy sampling instead of upper confidence bounds (UCB) that Lai (1987) has defined only in the one-parameter case. Auer, Cesa-Bianchi and Fischer (2002, Theorem 3) have shown that the ϵ-greedy sampling method introduced by Sutton and Barto (1998) can also attain the asymptotic lower bound for the regret[1].

S2 Proof of Theorem 1

Chan and Lai (2006, Section 5.2) consider the one-dimensional case of condition (C) but does not provide details[2] on how the procedure (which we have just described in the case $\inf_{\mathbf{x}} g(\mathbf{x}) > 0$) still yields under (C) "an asymptotically optimal selection procedure with expected total sampling cost of the order of $|\log \alpha|$". Since condition (C) is the counterpart of the assumption $\inf_{\mathbf{x}} g(\mathbf{x}) > 0$ in Theorem 2, the preceding proof of Theorem 2 has already provided these details even for the multiparameter exponential family, thereby proving Theorem 1.

We now explain the "empirical Bayes hyperparameter tuning" refor-

[1]The result stated in that paper is for time-varying ϵ_t. It is also applicable to time-invariant ϵ when t does not exceed a finite upper bound T, which is assumed in Theorems 1 and 2.

[2]Chan and Lai (2006) only mention the asymptotic lower bound for the regret in multi-armed bandit problem or exponential families and refer to Lai (1987) for the UCB rule that attains the bound.

mulation of Theorem 1 mentioned in the paragraph following the theorem. Assume that g satisfies condition (C) and that $\mathbf{Z}_t$ are independent with density from a multiparameter exponential family. Let $\boldsymbol{\lambda}_{\text{opt}}$ denote the optimal hyperparameter of a metaheuristic optimization algorithm to minimize $C_T(\boldsymbol{\lambda}) = \sum_{j=1}^{J} g\left(\mathbf{x}^* - \mathbf{x}_{t_j}(\boldsymbol{\lambda})\right)\tau_j$ for a particular problem or system, and let $\Pi_{\boldsymbol{\lambda}}$ be a prior distribution on $\boldsymbol{\lambda}_{\text{opt}}$ so that $B(\boldsymbol{\lambda})$ is the Bayes rule that minimizes the Bayes risk $\int EC_T(\boldsymbol{\lambda})\, d\Pi_{\boldsymbol{\lambda}}$. The Empirical Bayes (EB) approach to hyperparameter tuning uses empirical performance to choose $\boldsymbol{\lambda} \in \Lambda$ for the Bayes rule $B(\boldsymbol{\lambda})$. In the group sequential setting of Section 2.1, an efficient group sequential EB hyperparameter tuning procedure is given by sequence $\{\widehat{\boldsymbol{\lambda}}_1, \ldots, \widehat{\boldsymbol{\lambda}}_J\}$ of Theorem 1, which shows that the sequence has asymptotically minimal Bayes risk, of order $O(\log B)$ as $B \to \infty$, among all group sequential hyperparameter tuning procedures with Bayes risks of order $o(B^r)$ for any $r > 0$.

S3 Figures 1 and 2 for Examples 1 and 2 (Section 4.1)

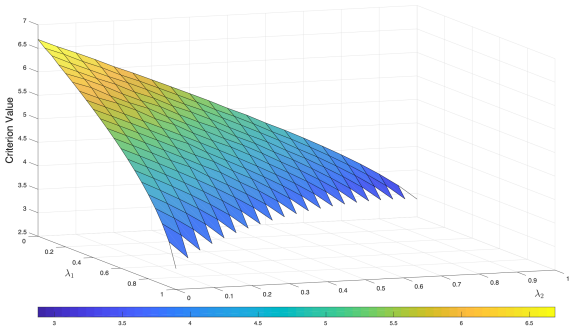

Figure 1: Pareto surface for compound optimality criterion (3.6), with colormap (using Matplotlib in MATLAB) explaining the colors of the surface.

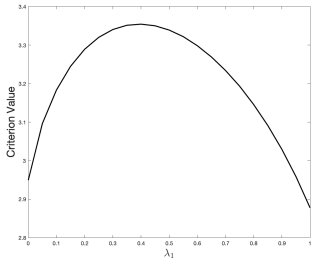

Figure 2: Pareto curve for (3.6) with $\lambda_3 = 0$ and known $\boldsymbol{\theta}$.

S4 Recent advances and background literature

We give below an overview, with additional references, of the recent advances in the theory and applications of adaptively tuned metaheuristic optimization algorithms, which are mentioned in Section 5. We use the

same alphabetic labeling of these advances as Section 5.

(a) Lai, Sklar, and Weissmueller (2020) describe the application of the efficient group sequential design methodology for early-phase dose-finding and efficacy testing trials to the Fast Real-time Assessment of Combination Therapies in Immuno-Oncology (FRACTION) platform trial recently "designed to rapidly evaluate new combinations of I-O agents and targeted therapies" by Bristol-Myers Squibb (Simonsen et al., 2018). A key feature of FRACTION is that "new combination regimens are added to the ongoing study as they become available", while "combination treatment arms in FRACTION studies that demonstrate futility will be terminated early, whereas those arms that meet early efficacy criteria will enroll additional patients" to obtain more precise estimates of the treatment effect or to test for other efficacy endpoints in Phase II/III registration trials. Moreover, "the safety, dose, schedule and preliminary antitumor activity of new combinations will be determined separately in prior Phase I studies." This is, therefore, in the spirit of efficient group sequential designs of early-phase trials in Section 4.2, which considers a single new combination therapy instead of multiple combination therapies in the "master protocol" of FRACTION. In fact, the computational advantages, for this task, of the adaptively tuned PSO algorithm are even more conspicuous for multiple combination ther-

apies than a single one as Lai, Xu, and Weismueller[3] have recently shown how the Phase I/II component of the group sequential design of early-phase dose-finding trials in Section 4.2 can be simultaneously computed by PSO for multiple combination therapies.

As pointed out by Woodcock and LaVange (2017) and FDA's 2018 Guidance for Industry on Master Protocols, a master protocol refers to an overarching clinical trial design that evaluates multiple hypotheses, with the objective to improve efficiency in the development of different interventions. While there were fewer than 10 master-protocol-guidance studies in the public domain by 2010, the subsequent decade saw rapid growth of master protocols in support of clinical development, leading to 83 clinical trials in the public domain that used master protocols. This rapid growth was catalyzed by the successful immuno-oncology (I-O) therapies inhibiting CTLA-4 (cytotoxic T lymphocyte-associated antigen 4) protein receptor (ipilimumab) and the PD-1 (programmed death-1) signaling pathway (pembrolizumab and nivolumab). Simonsen et al. (2018, p.260) point out that "beyond PD-1 and CTLA-4 blockade therapies, a rapidly growing number of novel I-O agents are in early development, supported by preclinical antitumor activity" involving other receptors and signaling pathways,

[3]Nikolas Weismueller is Associate Director of Advanced Analytics and Real World Data at Bristol-Myer Squibb.

and that "combination therapies may improve treatment outcomes relative to monotherapies because of their additive or synergistic effects". Basket trials, platform and umbrella trials are implementation structures of master protocols. Whereas basket trials investigate a single drug or a single combination therapy across multiple patient populations, umbrella and platform trials involve multiple therapies in one or multiple patient populations, with addition or removal of studies specified in the master protocol of a platform trial. FRACTION is a platform trial that has been motivated by "the success of combination therapy with nivolumab and ipilimumab (that was approved for treatment of metastatic melanoma, which) suggests that other combination approaches for modulating immunosuppression may also be applicable to other malignancies" including non-small cell lung cancer, renal cell carcinoma, gastric and esophageal cancers. "When additional combination therapies have sufficient safety data and scientific rationale (from prior Phase I trials) to enter FRACTION, they will be introduced via a new Sub-Protocol", which can also incorporate its own design features.

(b) We begin with additional references and further discussion on Phase II in Section 4.2 and its extension to a late-phase confirmatory trial via a seamless Phase II/III design. In this connection we also supplement the overview of Phase I in Section 4.1 and the second paragraph of Section 4.2

with Chapter 2 and Section 4.1 of Chen, Heyse and Lai (2018), abbreviated by CHL hereafter. In particular, besides summarizing rule-based and model-based Phase I designs, Section 4.1 of CHL also introduces threshold designs of Ji, Li, and Bekele (2007) that use model-based "toxicity probability intervals" to address the uncertainty in the MTD estimates from the observed toxicity outcomes, and extensions to combination therapies by Yin and Yuan (2009), Lee, Fan, and Lu (2017) and others. Chapter 2 of CHL gives an overview of pharmacokinetic-pharmacodynamic (PK-PD) models that quantify responses to a drug through (i) the time course of drug concentration in plasma or blood after drug administration during which the processes of absorption, distribution, metabolism and elimination are studied (PK), and (ii) the relationship between drug concentration at the effect site and therapeutic effects which can be efficacy or toxicity effects or both (PD). It also introduces modern machine learning methods to estimate a quantitative structure-toxicity relationship (QSTR) or quantitative structure-activity relationship (QSAR) for the prediction of toxicity or biological activity from the chemical attributes and/or physical properties of the drug molecules. We can now add adaptively tuned metaheuristic optimization algorithms to the machine learning/AI. Concerning additional references on Phase II designs in Section 4.2, Section 4.2 of CHL gives an

overview of the safety considerations for the design of Phase II and Phase III clinical trials and describes the REST (Rotavirus Efficacy and Safety Trial) of Merck's rotavirus vaccine and its statistical analysis by conditioning on rare adverse events (intussusceptions during infancy in the case of rotavirus vaccines); see Heyse et al. (2008) who "basically use a repeated significance test that terminates the study after n intussusceptions cases are observed" (p.108 of CHL). After hearing a presentation by Heyse and Jie Chen on their design based on sequential GLR (generalized likelihood ratio) statistics could be used to improve the REST design and analysis. This led to Shih et al. (2011) and Lai's decade-long collaboration with Chen and Heyse; see Chapter 5 of Bartroff, Lai and Shih (2013), whose Section 6.7 (Supplement 3) and Section 7.5 contain information and discussions of Phase II/III design of oncology trials. Moreover, in addition to BLN (Bartroff, Lai and Narasimhan, 2014) that Phase II in Section 4.2 has focused on, we should also refer to Dale (1986), Yin, Li, and Ji (2006) and Yuan and Yin (2011) whose works paved the way for the comprehensive development in Section 3.2 of BLN.

We next consider the issue of valid statistical inference from the data in adaptive confirmatory group sequential trials, e.g., those with master protocols or at the termination of a Phase II/III trial. Section 3.1 of Lai, Sklar,

and Weissmueller (2020) describes how the hybrid resampling method introduced by Chuang and Lai (1998, 2000) and subsequently extended by Lai and Li (2006), Lai, Shih, and Su (2009) can be used to analyze these confirmatory group sequential trials; see Section 7.3 and 7.4 of Bartroff, Lai, and Shih (2013). The methodology has been recently extended to include variable selection and multiple testing for high-dimensional covariates and change-point time series models by Lai, Choi, and Tsang (2019) and Dai and Tsang (2020). The last sentence of Section 1 remarks that some of the designs and analytic methods for oncology platform trials are also applicable to the recent adaptive platform trials for COVID-19 vaccine and drug development. We refer to Chen, Choi, and Lai (2020) who "begin with a description of the collaborative public-private partnership known as Accelerating Covid-19 Therapeutic Interventions and Vaccines (ACTIV) Initiatives" and the statistical, computational and commercial challenges in data sharing among the partner companies and government agencies, and then "describe the basic science of SARS-Cov-2" and "discuss some recent advances in adaptive confirmatory trial designs and valid statistical inference methods to ensure reproducible findings from such trials in a highly adaptive setting."

(c) Other metaheuristic algorithms include quantum PSO mentioned

in Section 1 and differential evolution (DE) which was recently used by Xu et al. (2019) to find high-dimensional D-optimal designs for logistic models. DE was proposed by Storn and Price (1997) as an enhancement of the Genetic Algorithm (GA) that dated back to John Holland and his students in the 1970s; see Goldberg (1989) and the R package DEoptim (Global Optimization by Differential Evolution), version 2.2.5 in CRAN. Both GA and DE have their origins in genetics and their operations involve *mutation, crossover*, and *selection*; see Xu et al. (2019, pp.7135-7136) who propose the following adaptive enhancement of mutation to be used in conjunction with a new crossover method called *multiple exponential recombination* (MER) for gradient-free high-dimensional optimization. The selection and mutation components are integrated into a "novelty-based mutation strategy" which selects a group of individuals that explore different and novel regions in the search space, and which "can balance exploration and exploitation at the early or medium stage of evolution", hence the name "novelty-based DE" (NovDE) of this DE enhancement, details of which are given in Xu et al. (2019, pp.7137-7138). In particular, the scaling factor F for mutation and crossover rate CR vary with the stage i of the evolution, thereby choosing adaptively the tuning parameter $\boldsymbol{\lambda} = (\mathrm{F}, \mathrm{CR})$ of NovDE which already uses a novelty-based mutation strategy for "exploration and exploitation",

similar to the multi-armed bandits in Section 2.2.

In his survey of the major developments during the past seven decades on stochastic approximation (SA) that was "introduced in 1951 to provide a new theoretical framework for root finding and optimization of a regression function in the then-nascent field of statistics", Lai (2020) describes in his Section 2.3 a general approach to the convergence proof of recursive stochastic algorithms via an "almost supermartingale" introduced by Robbins and Siegmund (1971) and later generalized to an "extended stochastic Liapunov function" by Lai (1989). His Section 3.3 shows how this approach can be applied to prove weak convergence of adaptive PSO, the non-adaptive version of which was proved earlier by Yuan and Yin (2015). The same approach can be used to prove almost sure convergence by using the stability bounds established by Tong et al. (2020). This general framework can clearly be applied to other metaheuristic optimization algorithms such as differential evolution by developing similar stability bounds.

(d) Multi-armed bandits with covariates, also called "contextual multi-armed bandits", arise in many fields of application in the current Big Data and Multi-Cloud era, ranging from personalized health and medicine to online personalized advertising and recommender systems. This has called for the extension of classical bandit theory in Section 2.2 to nonparametric

contextual bandits, and Kim, Lai, and Xu (2020) have recently provided a definitive extension. This extension leads to the far-reaching generalization of Theorems 1 and 2 from the exponential family to semiparametric families. The key underlying idea is to replace the UCB rule in classical multi-armed bandits by ϵ-greedy randomization and an arm elimination scheme. Details are given in Section 2.2 of Kim, Lai, and Xu (2020) whose Section 3 also extends the theory to high-dimensional covariates, for which recent advances in machine learning for recommender systems (e.g., Dai et al. (2019)) have enabled implementation of the theory.

Additional References

Bartroff, J., T. L. Lai, and M.-C. Shih (2013). *Sequential Experimentation in Clinical Trials: Design and Analysis*, Volume 298. Springer, New York.

Bechhofer, R. E., J. Kiefer, and M. Sobel (1968). *Sequential Identification and Ranking Procedures*, Volume 3. University of Chicago Press.

Chan, H. P. and T. L. Lai (2006). Sequential generalized likelihood ratios and adaptive treatment allocation for optimal sequential selection. *Sequential Analysis 25*(2), 179–201.

Chen, J., A. Choi, and T. L. Lai (2020). Challenges and opportunities in

designing clinical trials for covid-19 treatments. *Statistics in Biopharmaceutical Research*, to appear.

Chuang, C.-S. and T. L. Lai (1998). Resampling methods for confidence intervals in group sequential trials. *Biometrika 85*(2), 317–332.

Chuang, C.-S. and T. L. Lai (2000). Hybrid resampling methods for confidence intervals. *Statistica Sinica 10*, 1–50.

Dai, B., J. Wang, X. Shen, and A. Qu (2019). Smooth neighborhood recommender systems. *The Journal of Machine Learning Research 20*(1), 589–612.

Dai, W. and K. W. Tsang (2020). Hybrid resampling confidence intervals for change-point or stationary high-dimensional stochastic regression models. *Statistica Sinica*, to appear.

Dale, J. R. (1986). Global cross-ratio models for bivariate, discrete, ordered responses. *Biometrics*, 909–917.

FDA (2018). Master protocols: efficient clinical trial design strategies to expedite development of oncology drugs and biologics. *Draft Guidance for Industry. Rockville, MD: Food and Drug Administration*.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA.

Heyse, J. F., B. J. Kuter, M. J. Dallas, and P. Heaton (2008). Evaluating the safety of a rotavirus vaccine: the rest of the story. *Clinical Trials 5*(2), 131–139.

Ji, Y., Y. Li, and B. N. Bekele (2007). Dose-finding in phase I clinical trials based on toxicity probability intervals. *Clinical Trials 4*(3), 235–244.

Kim, D., T. Lai, and H. Xu (2020). Multi-armed bandits with covariates: Theory and applications. *Statistica Sinica*, to appear.

Lai, T. L. (1989). Extended stochastic Lyapunov functions and recursive algorithms in linear stochastic systems. In *Stochastic Differential Systems*, pp. 206–220. Springer.

Lai, T. L. (2020). Stochastic approximation: From statistical origin to big-data, multidisciplinary applications. *Statistical Science*, to appear.

Lai, T. L., A. Choi, and K. W. Tsang (2019). Statistical science in information technology and precision medicine. *Annals of Mathematical Sciences and Applications 4*(2), 413–438.

Lai, T. L. and W. Li (2006). Confidence intervals in group sequential trials with random group sizes and applications to survival analysis. *Biometrika 93*(3), 641–654.

Lai, T. L., M.-C. Shih, and Z. Su (2009). Tests and confidence intervals for secondary endpoints in sequential clinical trials. *Biometrika 96*(4), 903–915.

Lai, T. L., M. Sklar, and N. T. Weissmueller (2020). Novel clinical trial designs and statistical methods in the era of precision medicine. *Statistics in Biopharmaceutical Research* (in press).

Lee, B. L., S. K. Fan, and Y. Lu (2017). A curve-free bayesian decision-theoretic design for two-agent phase I trials. *Journal of Biopharmaceutical Statistics 27*(1), 34–43.

Robbins, H. and D. Siegmund (1971). A convergence theorem for non-negative almost supermartingales and some applications. In *Optimizing Methods in Statistics.*, pp. 233–257. Academic Press.

Shih, M.-C., T. L. Lai, J. F. Heyse, and J. Chen (2011). Sequential generalized likelihood ratio tests for vaccine safety evaluation. *Statistics in Medicine 29*(26), 2698–2708.

Simonsen, K. L., P. M. Fracasso, S. H. Bernstein, M. Wind-Rotolo, M. Gupta, A. Comprelli, T. P. Reilly, and J. Cassidy (2018). The fast real-time assessment of combination therapies in immuno-oncology (fraction)

program: innovative, high-throughput clinical screening of immunotherapies. *European Journal of Cancer 103*, 259–266.

Storn, R. and K. Price (1997). Differential evolution a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization 11*, 341–359.

Sutton, R. S. and A. G. Barto (1998). *Introduction to Reinforcement Learning*, Volume 135. MIT Press.

Tong, X., K. P. Choi, T. L. Lai, and W. K. Wong (2020). Stability bounds and almost sure convergence of improved particle swarm optimization methods. *Research in Mathematical Sciences*, to appear.

Woodcock, J. and L. M. LaVange (2017). Master protocols to study multiple therapies, multiple diseases, or both. *New England Journal of Medicine 377*(1), 62–70.

Xu, W., W. K. Wong, K. C. Tan, and J.-X. Xu (2019). Finding high-dimensional d-optimal designs for logistic models via differential evolution. *IEEE Access 7*, 7133–7146.

Yin, G., Y. Li, and Y. Ji (2006). Bayesian dose-finding in phase I/II clinical trials using toxicity and efficacy odds ratios. *Biometrics 62*(3), 777–787.

Yin, G. and Y. Yuan (2009). Bayesian dose finding in oncology for drug combinations by copula regression. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 58*(2), 211–224.

Yuan, Q. and G. Yin (2015). Analyzing convergence and rates of convergence of particle swarm optimization algorithms using stochastic approximation methods. *IEEE Transactions on Automatic Control 60*(7), 1760–1773.

Yuan, Y. and G. Yin (2011). Bayesian phase I/II adaptively randomized oncology trials with combined drugs. *The Annals of Applied Statistics 5*(2A), 924–942.