A System-Level Analysis of Conference Peer Review

YICHI ZHANG, University of Michigan
FANG-YI YU, Harvard University
GRANT SCHOENEBECK, University of Michigan
DAVID KEMPE, University of Southern California

We undertake a system-level analysis of the conference peer review process. The process involves three constituencies with different objectives: authors want their papers accepted at prestigious venues (and quickly), conferences want to present a program with many high-quality and few low-quality papers, and reviewers want to avoid being overburdened by reviews. These objectives are far from aligned; the key obstacle is that the evaluation of the merits of a submission (both by the authors and the reviewers) is inherently noisy. Over the years, conferences have experimented with numerous policies and innovations to navigate the tradeoffs. These experiments include setting various bars for acceptance, varying the number of reviews per submission, requiring prior reviews to be included with resubmissions, and others. The purpose of the present work is to investigate, both analytically and using agent-based simulations, how well various policies work, and more importantly, why they do or do not work.

We model the conference-author interactions as a Stackelberg game in which a prestigious conference commits to a threshold acceptance policy which will be applied to the (noisy) reviews of each submitted paper; the authors best-respond by submitting or not submitting to the conference, the alternative being a "sure accept" (such as arXiv or a lightly refereed venue). Our findings include:

- observing that the conference should typically set a higher acceptance threshold than the actual desired quality, which we call the *resubmission gap* and quantify in terms of various parameters.
- observing that the reviewing load is heavily driven by resubmissions of borderline papers — therefore, a judicious choice of acceptance threshold may lead to fewer reviews while incurring an acceptable loss in quality.
- observing that depending on the paper quality distribution, stricter reviewing may lead to higher or lower acceptance rates — the former is the result of self selection by the authors.
- finding that a relatively small increase in review *quality* or in self assessment by the authors is much more effective for conference quality control (without a large increase in review burden) than increases in the *quantity* of reviews per paper.
- showing that keeping track of past reviews of papers can help reduce the review burden without a decrease in conference quality.

For robustness, we consider different models of paper quality and learn some of the parameters from real data.

CCS Concepts: • **Theory of computation** → *Algorithmic mechanism design*; *Algorithmic game theory*.

Additional Key Words and Phrases: Peer review, Stackelberg game

YZ and GS were supported by NSF award number # 2007256. FYY was supported by NSF IIS-2007887. DK was supported in part by ARO grant W911NF1810208.

EC '22, July 11–15, 2022, Boulder, CO, USA.

ACM ISBN 978-1-4503-9150-4/22/07…$15.00
https://doi.org/10.1145/3490486.3538235

ACM Reference Format:
Yichi Zhang, Fang-Yi Yu, Grant Schoenebeck, and David Kempe. 2022. A System-Level Analysis of Conference Peer Review. In *Proceedings of the 23rd ACM Conference on Economics and Computation (EC '22), July 11–15, 2022, Boulder, CO, USA.* ACM, New York, NY, USA, 40 pages. https://doi.org/10.1145/3490486.3538235

1 INTRODUCTION

Conferences play an important role in the publication and scientific dissemination process in computer science. They aim to provide attendees with access to high-quality and recent research results[1], and authors view the ability to publish and share their recent results at high-quality venues as conferring scientific credibility and thus status. In order to do so, conferences rely on significant volunteer work from the community, most notably in reviewing large numbers of submitted papers to evaluate their scientific merit.

While the conference publication process seems to have served the community fairly well overall — in particular enabling a high speed of dissemination of scientific results — different members of the community often also see significant room for improvement: authors often feel that their submitted papers are not evaluated sufficiently competently[2], while conference attendees sometimes find the program diluted with less interesting work. Attempts have been made by conferences to mitigate such concerns, one of which is to increase the number of reviews assigned to each submission.[3] However, this approach increases the review burden; indeed, many members of the scientific community now feel overloaded with conference reviewing requests. The approach has the potential to lead to a vicious cycle: the substantial increase in peer review workload[4] may lead to lower-quality reviews, in turn leading to more resubmissions of the same papers and thus a higher review burden.

Our high-level goal is to understand how the parameters of the system, together with the conference's review policy decisions, affect the tradeoff between the conference quality and the review burden on the community. Given a fixed number and quality distribution of submissions, the conference's quality is increased by accepting more good and fewer bad papers. Distinguishing between good and bad papers more accurately requires more reviews, along with enticing self selection by authors.

We model the conference review process as a Stackelberg game between a top conference and the (homogeneous and strategic) authors (described in detail in Section 2). The authors' papers have qualities (positive or negative) drawn from a commonly known distribution over a (finite or infinite) set; however, both the conference's reviewers and (in some of our analysis) the authors themselves only obtain imperfect signals about the quality of a paper. The conference's quality is the sum of qualities of the accepted papers, normalized by the total number of papers. The conference commits to a review and acceptance policy, which prescribes how many independent reviewers are assigned to each paper, and what the criteria for acceptance are. In Appendix D, we also study policies with "institutional memory," including policies limiting the number of times a paper can be resubmitted, and policies requiring past reviews to be included with resubmissions.

In response to the conference's policies, the authors face a binary decision in each round: whether to submit to the top conference (which leads to some positive utility if the paper is accepted, but zero utility if the paper is rejected), or submit to a safe choice (such as a second-tier conference or arXiv) and get a guaranteed smaller positive utility. The utility of acceptance to either venue is exponentially time-discounted in the number of resubmission rounds, modeling that authors

[1] in addition to networking opportunities
[2] and of course always with reviewers erring on the negative side
[3] Assigning five reviewers to one paper is no longer rare for top computer science conferences.
[4] also exacerbated by the growth of the research community

prefer timely acceptance of their result. Authors best-respond, i.e., choose a utility-maximizing action, based on their private (potentially noisy) signal about their paper's quality as well as the historical reviews collected when the paper was rejected in previous rounds. We primarily focus on the following high-level questions:

- The fact that rejected papers can be resubmitted means that weaker papers may be eventually accepted if authors are patient enough. What is the impact of resubmissions on the gap between a conference's acceptance threshold and the de facto distribution of paper qualities at the conference?
- What is the range of Pareto optimal review and acceptance policies with regard to the tradeoff between conference quality and review burden? How is it impacted by conference prestige, authors' patience, and/or review quality?
- How does the conference's acceptance policy affect its acceptance rate? Note that this is non-obvious, as a strict policy may lead to a lot of self selection on the part of the authors, and thus to a *higher* acceptance rate.
- How does the number of reviews per round affect the overall review burden? What is the tradeoff between review quality and quantity, i.e., how many noisy reviews approximate one high-quality review?
- How helpful is institutional memory? Can the conference significantly lower the review burden or increase quality by limiting the number of resubmissions, treating resubmitted papers differently, or requiring past reviews to be included?

To ensure robustness of our results, we investigate these questions under different models regarding (1) the quality distributions of papers and distributions of review noise, and (2) the information authors have about the quality of their own work.

For (1), we consider three models: (a) a model with continuous paper qualities (e.g., drawn from a Gaussian distribution), where reviews are also continuous with additive noise, (b) a class of categorical models with several discrete paper qualities and a discrete scale of review scores, and (c) a model with binary paper qualities (good/bad) and reviews (accept/reject recommendations). The advantage of (a) and (c) is that they are characterized by a small number of meaningful parameters, making it possible to systematically explore the dependence of outcomes on these parameters. On the other hand, (b) allows for more realistic modeling of real-world conferences, and we estimate the parameters from publicly available review data for the ICLR 2020 and 2021 conferences [1, 2].

For (2), we consider authors with perfect information about their paper's intrinsic quality, and authors who themselves only receive noisy signals. The advantage of the former model is that authors do not learn new information from reviews, so their decision is either to submit their paper until it is accepted, or to immediately submit to the sure bet option. In turn, this simplicity of best responses allows us to prove theoretical results about the outcomes. On the other hand, this behavior is also somewhat unrealistic, motivating the study of a model in which authors themselves have noisy signals, and update their beliefs about the paper's quality based on the reviews. The resulting Bayesian reasoning makes the model too complex for a theoretical analysis, so we investigate it using agent-based modeling (ABM) and simulations.

We emphasize results in the model which portrays them most clearly. We hope that this helps the reader quickly grasp the underlying intuition. However, we also typically test the robustness of the results in the remaining models. In particular, many of our results are shown theoretically in simple models, but are shown to continue to hold (at least qualitatively) with simulations for the more complex models.

While the models are necessarily a simplification, and one should therefore be very cautious of directly basing concrete decisions on the results, we hope that our theoretical results (and simulations) can steer the discussion, uncover parameters to focus on, and inform decision makers.

1.1 Summary of Results

Noiseless Authors: Theoretical Results. We first assume that authors know the quality of their manuscript, but reviewers only obtain noisy signals of the quality.

It is sometimes argued that conferences should accept every paper that is "above the bar." Because every submitted paper will be resubmitted until accepted (since the authors learn no new information from the reviews), the conference's *acceptance threshold* induces a *de facto threshold*: a manuscript quality above which every paper will be accepted and below which no manuscript will be submitted. The acceptance threshold and de facto threshold are typically different, sometimes significantly so (in particular when the conference is very attractive to authors, the authors are very patient, or the noise of reviews is large) — we call the difference the *resubmission gap*. If a conference naively sets an acceptance threshold identical to their ideal de facto threshold, they will ultimately accept substandard papers, in rounds in which enough reviewers had noisy reviews that ended up too high. Instead, the conference should select an acceptance threshold higher than the desired threshold of conference acceptances. However, the resulting "optimal" is then counter-intuitive: every paper that is submitted is eventually accepted, yet each round many of the papers are rejected. We exactly characterize the resubmission gap in several settings.

Second, we consider the tradeoff between the conference's quality and the review burden (the total number of reviews for each paper throughout its resubmission process), focusing on the Pareto frontier. At a high level, the review burden tracks the number of papers just above the de facto threshold, because such papers typically require numerous resubmissions before acceptance. A conference may be able to significantly reduce the long-term review burden by choosing the threshold such that there are fewer borderline papers near the resulting de facto threshold.

Third, we show that if the prestige of the conference increases, the patience of authors increases, or the noise of the reviews increases (in a certain technical sense), then the tradeoff between the conference's quality and the review burden becomes strictly worse. Thus, one cause of more reviews may be the high prestige placed on certain venues. This also warns that policies which effect these items may have unintended consequences.

Fourth, we study what impacts the acceptance rate of the conference. We show that as the noise of reviewers' signals shrinks, the relation between the acceptance threshold and acceptance rate is related to the *hazard rate* of the prior distribution of papers. If the distribution of paper qualities has a monotone (increasing) hazard rate, then as the acceptance threshold increases, the acceptance rate decreases. However, for priors with thick tails where the hazard rate may decrease, a stricter acceptance threshold often leads to a larger acceptance rate. We illustrate the robustness of these observations using simulations.

Noiseless Authors and Real-World Parameters. We further investigate the above phenomena in a more realistic setting based on categorical data. One common attempt at improving a conference's quality is to solicit more reviews per submission which, as believed, can help distinguish the good papers from the bad ones. Thus, we next (in Appendices B.4 and B.5) study how the number of solicited reviews for each submission impacts the conference's quality and the average total number of reviews for each paper. (Recall that each paper may be submitted multiple times.) We find that in our models, increasing the number of solicited reviews per paper beyond three rarely leads to better outcomes. This is because a small number of solicited reviews can be optimal with a

carefully chosen acceptance threshold. In contrast, we find that increasing the *quality* of reviews often substantially improves the Pareto frontier.

Authors with Noisy Signals. We continue to investigate the robustness of the above results in a setting where authors do not exactly know the quality of their paper and can learn about it from reviews. In this context, the Pareto frontier for conference quality vs. review burden is significantly worse than with perfectly appraised authors — the reason is that the latter can be compelled to self-select with carefully chosen acceptance thresholds. However, it is not clear if such gains are realizable in practice. Perhaps, new discipline norms of learning the quality of one's paper prior to submission (for example, by sharing early manuscripts with colleagues for feedback) could provide authors with accurate quality signals. Or perhaps, in practice, overcoming authors' unfounded admiration of their own work is not possible.

Memory in the System. Another popular type of proposal is to give the system more memory, either by limiting the number of resubmissions of the same manuscript or by reusing reviews.[5] Our results here (in Appendix D) show that the main effect of having memory within the review process is that the conference can reduce the review burden while preserving the same (or slightly better) conference quality. However, such an effect can be marginal in some cases; thus, it is not clear whether it is worthwhile to implement these policies broadly. It should be noted, though, that our analysis here is rather preliminary, and only carried out for the binary model.

Our models are primarily designed to investigate the tradeoff between conference quality and the reviewing burden for the community. They necessarily abstract away several other aspects of the conference submission ecosystem which would also be worth investigating, most importantly the authors' utility. These limitations are discussed in more depth in Section 5.2

Due to space limitation, our results on *Noiseless Authors and Real-World Parameters*, *Authors with Noisy Signals* and *Memory in the System* are presented in Appendices B, C, and D, respectively.

1.2 Related Work

Not surprisingly given the importance of peer review in science, several attempts have been made by different research fields to simulate, understand, and improve the process. When considering the systems level, agent-based models have been one of the techniques of choice. The review article by Feliciani et al. [12] gives a fairly comprehensive summary of this line of work. It suggests several general themes to focus on in models: editorial strategies, matching submissions with reviewers, decision making, biases and calibration, and comparisons of alternative peer review systems.

Among the more prominent works using ABMs are those by Kovanis et al. [17, 18]. Kovanis et al. [17] propose a model for a holistic study of the scientific publication ecosystem — this model includes the acquisition of resources (such as status) by authors, which can be leveraged into future papers. [18] builds on these models and implementations to evaluate several alternative systems for peer review. These models and results differ from ours in several key dimensions: the authors are not strategic, they do not focus on fine-grained policies by journals (or in our case, conferences), and due to the holistic nature and complexity of the model, the model is only amenable to simulation, but not analytically tractable.

Two papers by Bianchi et al. [5] and Squazzoni and Gandelli [26] also use agent-based modeling approaches. They particularly focus on the fact that researchers must decide how to divide their time between writing and reviewing papers, and investigate (experimentally) the impact of various policies on the efficiency of peer review. Similarly, Thurner and Hanel [29] and D'Andrea and

[5] Another reason for requiring the inclusion of prior reviews — not modeled here — is that it lets the conference ascertain that specific concerns from earlier versions have been addressed.

O'Dwyer [10] use agent-based models to investigate a specific aspect of peer review, namely, selfish behavior on the part of referees, who may not have incentives to see other strong work published.

Allesina [3] also uses agent-based modeling, in this case to understand the impact of different high-level approaches (editorial desk rejects, bidding on papers, etc.) on the overall reviewing load. Roebber and Schultz [22] use agent-based modeling to evaluate strategies for program officers of funding agencies. One of their findings is similar to ours: that requiring unanimous support for accepting a proposal (i.e., setting a high threshold) can discourage authors from submitting many proposals, thus lowering the review burden.

A more analytical approach is taken by Smith and Wilson [25]. Here, the authors are also interested in the impact of self selection on the acceptance rate of a journal (or university). They study a system with multiple journals or universities announcing different thresholds in the presence of noisy reviews. Due to their motivation, their model does not appear to account for resubmissions, thus differing from our work in a key aspect.

Several other works do more basic theoretical analysis of the impact of conference policies. In particular, they focus on the false positives (accepting bad papers) and false negatives (rejecting good papers) arising as a function of the number of reviewers and their individual qualities. Based on such calculations, Herron [14] suggests that obtaining a large number of low-quality reviews may be better than a small number of expert reviews. Neff and Olden [20] focus in particular on the role of desk rejects by an expert editor.

In response to concerns by authors about the evaluation of their work when submitted to conferences or journals, the scientific community in general, and CS community in particular, has engaged in significant self evaluation efforts. Many of these have focused on the quality and consistency of reviews provided to conferences [6, 11, 24]. Cole et al. [7] and Tran et al. [30] study the reproducibility and randomness in review scores and acceptance decisions. Furthermore, the community has experimented with (or at least suggested) different formats, including increasing the number of reviews per paper, multi-level or multi-stage evaluation processes, having reviews from past submissions follow a paper upon resubmission, and many others [15, 21, 23, 24]. Many of these approaches appear primarily driven by concern for authors and their desire for accurate evaluation of their submission, though some of them are also part of our evaluation.

Other attempts that try to mitigate the overwhelming demand for reviewers consider solutions based on mechanism design. Srinivasan and Morgenstern [27] combine a bidding system and peer prediction to simultaneously incentivize high-quality reviews and high-quality submissions. Su [28] designs a mechanism that elicits ranking information truthfully from the authors, which is proven to empirically benefit the conference's quality.

In our idealized model, we assume that all reviews are i.i.d. Naturally, this is a simpification. In reality, one of the difficulties faced by conferences and journals is how to aggregate the scores from reviewers with possibly very different scales or expectations. Indeed, such aggregation is a well-known fundamental problem in statistics [4, 8, 31].

Peer review can also be viewed through the lens of a principal-agent problem: the principal decides on the review process and the acceptance rule, and the agents respond. Besides peer review, related applications include admitting college students [16] and recruiting faculty [32], or endorsing a product [13, 19]. In particular, Gill and Sgroi [13], Lerner and Tirole [19] study models in which the agent (such as an author) can choose among venues one that maximizes the expected utility, as determined by the chance of success and the prestige.

2 MODEL

We consider a process of an *author* (or group of authors) submitting a paper[6] to a prestigious *conference*[7]. We model the submission-reviewing process as a multi-round game: in each round, the author decides to submit the paper either to the prestigious conference or an undiscriminating ("sure bet") conference[8] that will always accept. Upon submission, the prestigious conference will send the paper out for review and, based upon the reviews, decide to accept or reject. If the paper is rejected, the author sees the reviews, and faces the same decision problem in the next round.

In the game, two main agents actually make decisions: the author of the paper and the prestigious conference. (The "sure bet" conference simply accepts all papers, and the reviewers simply provide reviews.) Whenever we refer to "the conference" as a decision maker, we therefore always mean the prestigious conference.

Each paper has a quality Q drawn i.i.d. from a commonly known prior paper quality distribution $\boldsymbol{p}$ over the set of possible qualities $\boldsymbol{Q} \subseteq \mathbb{R}$; larger qualities correspond to better papers. (The assumption that $\boldsymbol{p}$ is commonly known is not essential — it only matters that the conference (e.g., PC chair) know the distribution.) Without loss of generality, there exist both negative and positive values in $\boldsymbol{Q}$; otherwise, the conference would simply accept/reject all papers without review. When the set of qualities is discrete, we write $p_q = \text{Prob}[Q = q]$; when it is continuous, we use $p(q)$ to denote the density of $\boldsymbol{p}$ at q. Because all papers' qualities are drawn from the same distribution, we will not need to reference a specific paper in our notation.

For each submission, the conference cannot observe its true quality, but will solicit some number, m, of reviews. Each review S_j, for $j = 1, \ldots, m$, is a random variable drawn i.i.d. from a distribution $\boldsymbol{\beta}_q$ where q is the paper's quality. The outcome of the jth review is $s_j \in \Sigma \subseteq \mathbb{R}$ where Σ is the set of possible review scores and a higher score denotes a more positive review. Thus, the reviews are independent conditioned on the paper's quality. We write $\boldsymbol{s}$ for the vector of the m reviews. We let $U(\boldsymbol{s}) = \mathbb{E}_Q\left[Q \mid \boldsymbol{s}, \boldsymbol{p}, \boldsymbol{\beta}\right]$ denote the expected quality of a paper conditioned on the reviews $\boldsymbol{s}$.

We assume the reviews to be *informative* about the paper's quality, in the sense that whenever $q < q'$, the corresponding distributions satisfy that $\text{Prob}_{s\sim\boldsymbol{\beta}_q}[s \geq x] < \text{Prob}_{s\sim\boldsymbol{\beta}_{q'}}[s \geq x]$ for all $x \in (\inf\Sigma, \sup\Sigma)$. That is, the distribution of review scores induced by a higher-quality paper first-order stochastically dominates that induced by a lower-quality paper. Furthermore, when $\boldsymbol{Q}$ and Σ are continuous, we assume that for all x, $\text{Prob}_{s\sim\boldsymbol{\beta}_q}[s \geq x]$ is a continuous function of q.

For some of our results, we assume that the authors perfectly observe Q, the paper's quality. This model has the advantage of being analytically tractable, because the authors do not learn new information from the reviews. We call such authors *noiseless*. For other results, we consider *noisy* authors, who only assess their papers' qualities approximately. In this case, we assume that authors have noisy signals $S^{(a)}$, which, similar to the conference's signals, are drawn according to some informative (and continuous, if $\boldsymbol{Q}$ and Σ are continuous) distribution $\boldsymbol{\alpha}_q$ for a paper of quality q. The author's signal is independent of the conference's signals conditioned on Q. Noisy authors will update their beliefs about their papers' qualities based on the reviews in a Bayesian way.

We primarily focus on two settings for quality and signals:

> **Categorical Model:** Here, both $\boldsymbol{Q} \subset \mathbb{R}$ and $\Sigma \subset \mathbb{R}$ are finite (but ordered by their natural order on $\mathbb{R}$). Such a model is equivalent to a modified version of the Dawid-Skene model [9] where there is an additional ordering categorical structure.

[6]The analysis focuses on the process for one paper. By considering the (independent) processes for multiple papers, we obtain a systems-level view. This is discussed in more detail in the section on the review burden, below.

[7]This single conference can refer to multiple "equivalent" conferences, e.g., STOC/FOCS.

[8]Equivalently, we can think of the undiscriminating conference as a preprint site like arXiv. Another viewpoint of this simplification is that the papers under consideration are of sufficient quality that the sure bet will always accept them.

Continuous Model: In the continuous model, we assume that both the paper quality and the review signals (conditioned on the paper quality) are drawn from a continuous distribution. Furthermore, the reviewers' signals are obtained by adding to the true quality Q a noise term drawn from a distribution $F^{(r)}$ which is *independent* of Q and has zero mean[9]. In other words, the cdf of the distribution of reviewer signals conditioned on a quality $Q = q$ is $F^{(r)}(x - q)$. We assume that $F^{(r)}$ is invertible.

2.1 Conference Acceptance Policy and Quality

The conference's lever of control is its acceptance policy. We primarily focus on *memoryless* acceptance policies. Under a memoryless acceptance policy, 1) the author can submit a paper an unlimited number of times; 2) in each round t in which the author (re-)submits the paper, the conference's decision depends only on the reviews obtained in round t; and 3) the same number of reviews m and decision rule is used in every round. In other words, each submitted paper is treated as a fresh paper. For that reason, we typically omit the round t from the notation when discussing memoryless policies. We discuss alternatives to memoryless policies in Appendix D, in particular, limiting the number of times a paper can be submitted, and having old reviews follow a resubmission. However, except for these sections, unless stated otherwise, all acceptance policies are memoryless.

A memoryless acceptance policy is characterized by a function $\phi : \Sigma^m \to [0, 1]$ which determines the probability with which each combination of review signals leads to a paper's acceptance. The conference's strategy is a pair m and ϕ. We assume that acceptance policies are both monotone — if reviews improve so does the probability of acceptance — and anonymous — the probability of acceptance does not depend on the order of the reviews. A particularly natural class of anonymous and monotone acceptance policies prescribe a conditional expected quality threshold.

Definition 2.1. A *threshold acceptance policy* ϕ_τ with threshold $\tau \in \mathbb{R} \cup \{-\infty, +\infty\}$ accepts a paper with reviews $\boldsymbol{s}$ when $U(\boldsymbol{s}) > \tau$ and rejects the paper when $U(\boldsymbol{s}) < \tau$.

Notice that when $\text{Prob}_s[U(\boldsymbol{s}) = \tau] > 0$, there may be multiple threshold acceptance policies with threshold τ, differing in the probability with which papers with conditional expected quality exactly τ are accepted or rejected. In general, we denote a threshold acceptance policy by $\phi_{\tau,r}$, where τ is the threshold and r is the acceptance probability of a paper with quality τ. However, when considering the continuous model, because the probability measure of this knife-edge event is zero, we simply omit the subscript r and use ϕ_τ to denote the threshold acceptance policy.

Once an acceptance policy ϕ is fixed, the probability of a submitted paper being accepted in a particular round is only a function of its underlying quality q. We denote this probability as $P_{\text{acc}}(\phi, q)$. We observe that when ϕ is monotone, $P_{\text{acc}}(\phi, q)$ is strictly monotone in q.

Proposition 2.2. *Let ϕ be a monotone acceptance rule. Then, $P_{acc}(\phi, q)$ is strictly monotone in q.*

The proof of this proposition, along with all other omitted proofs, can be found in Appendix A.

In round t of (re-)submission, we suppose that the conference faces a (the same) number K of new papers as well as the previously rejected papers from the past $t - 1$ rounds. We define the *quality* $U^{(c)}$ of the conference as the expected value of the sum of all the accepted papers' qualities in round t, normalized by K and taken in the limit of t. As typical in Stackelberg games, we assume that the authors best-respond and, if there are multiple best responses, break ties so as to help the conference. Note that when $t \to \infty$, the expected number of submissions in round t that have previously been submitted ℓ times converges for any $\ell \in \mathbb{N}$.

[9]The zero-mean assumption is without loss of generality so long as the noise distribution is independent of the quality Q, as any (known) bias could be subtracted out from the reviews.

2.2 Author Utility and Decisions

In terms of timing, first, the conference announces its review and acceptance policy; subsequently, in each round t, the author decides whether to submit the paper to the conference or the "sure bet" option. The game ends when the paper is accepted at the conference or at the "sure bet" option.

Authors are characterized by two parameters: their time discount factor η, capturing how patient they are, and the prestige $V > 1$ they ascribe to the conference. (We normalize the value of the sure bet option to 1.) When the paper is accepted at the conference in round t, the author's utility is therefore $U^{(a)} = \eta^{t-1} \cdot V$; when the paper is accepted at the sure bet option in round t, the author's utility is η^{t-1}. The η^{t-1} term encodes exponential time discounting and models that authors would like their work to be published in a timely manner. Besides the utility loss due to time discounting, rejection does not cause additional cost for the author.

The author's decisions depend on all available information, i.e., her own (possibly perfect) private signal $S^{(a)}$ as well as all the reviews she has received for previous submissions. We assume that the author is rational and Bayesian, so her decisions are based on posterior quality distributions taking into account all available information. She will submit to the conference in round t if and only if her expected utility from doing so (over all future time steps $t' \geq t$) exceeds her expected utility from the sure bet (which is exactly η^{t-1} at the point she is making the decision). Notice that unless the author obtains a perfect signal, the reviews she obtains (in addition to her own signal $S^{(a)}$) change her posterior conditional probability over the paper's quality, which in turn changes her belief of the probability distribution of future reviews.

Definition 2.3. The model in which authors have perfect information about their papers' qualities, and papers may be resubmitted an unlimited number of times is called the model of *noiseless authors with unlimited resubmissions*.

Under the model of noiseless authors with unlimited resubmissions, a theoretical analysis becomes more tractable. This is because authors' beliefs of their papers' qualities will not be updated based on the reviews. As a result, the papers that are submitted in the first round will be repeatedly resubmitted until acceptance. Consequently, the quality of the conference depends entirely on the authors' self-selection.

2.3 Review Burden and Tradeoffs

As we discussed in the introduction, we are primarily interested in the tradeoff between the conference's quality and the review burden, which is captured by the number of requested reviews imposed on the community.[10] We call this the QB-tradeoff. We denote the expected number of reviews of a paper by R, which will be called the *review burden*. To be precise, the conference's policy, along with the author's best response, determines a probability distribution $\boldsymbol{b}$, where $b_t = \text{Prob}[\text{The paper is submitted at least } t \text{ times}]$. Then, $R = m \cdot \sum_{t=1}^{\infty} b_t$. In the limit as the round becomes large, the reviewing load is spread out evenly over rounds. Therefore, with K new submissions in each round, the review burden on the community is KR.

We say that a conference's policy with $U^{(c)}$ and R *dominates* another one with $\hat{U}^{(c)}$ and $\hat{R}$ if it has a higher (or equal) expected utility, $U^{(c)} \geq \hat{U}^{(c)}$, the number of reviews is smaller (or equal), $R \leq \hat{R}$, and at least one of the inequalities is strict. Given a set of policies, a policy is *Pareto optimal* if it is not dominated by any other policy in the set. We say that one QB-tradeoff curve C *dominates* another QB-tradeoff curve C' if for any point on the QB-tradeoff C' that does not correspond to accepting all papers nor rejecting all papers, there exists a point on the QB-tradeoff curve C that

[10] This downplays the utility of the author, which is discussed in Section 5.2.

dominates it. Note that this implies that no point on the Pareto frontier of the QB-tradeoff curve C is dominated by any of the points on the QB-tradeoff curve C'.

3 NOISELESS AUTHORS: THRESHOLDS AND RESUBMISSION GAPS

We first focus on the case of noiseless authors. Recall that in this setting, because the author will not update her belief about the paper's quality, she will either submit to the conference until the paper is accepted, or immediately submit to the sure bet option. In turn, this allows us to obtain analytical expressions for the conference's utility and optimal strategy.

In this section, we focus on the *resubmission gap*: the difference between the threshold the conference sets for acceptance and the actual threshold of accepted papers, taking into account the resubmission of previously rejected papers. An analysis of the resubmission gap is of interest in its own right (since it may crucially inform conference acceptance policies), and also serves as the foundation of our further investigation of tradeoffs in Section 4.

3.1 The Author's Best Response: De Facto Thresholds

When the conference's acceptance policy ϕ is fixed, from the author's perspective, it induces an acceptance probability $P_{\text{acc}}(\tau, q)$ for each paper quality q. By Proposition 2.2, this probability is strictly monotone in q. We first show that as a result, the author's best response is characterized by some θ such that the author submits the paper (and resubmits until accepted) if and only if $q \geq \theta$.

For the statement and proof of the following proposition, and most others later, we capture the appeal of submitting to a conference by the attractiveness factor ρ (where smaller values are better), based on its value V and the discount factor η:

$$\rho := \frac{1-\eta}{V-\eta}. \tag{1}$$

Because our results depend on η and V only through ρ, we observe that they are in a sense "interchangeable," albeit not linearly. That is, an increase in author patience (η) is tantamount to a (different) increase in conference prestige, as far as author behavior is concerned.

Proposition 3.1. *Consider a memoryless conference with acceptance policy ϕ. Let the authors be noiseless, with value for acceptance $V > 1$ and discount factor $\eta \in (0,1)$. Let $\theta = \min\{q \mid P_{acc}(\phi, q) \geq \rho\}$. A rational author will submit a paper of quality $Q = q \geq \theta$ to the prestigious conference until it is accepted, and if $Q = q < \theta$, will send the paper to the sure bet.*

The proposition implies that every monotone conference acceptance rule induces a corresponding submission threshold θ such that w.l.o.g., noiseless authors submit all papers with quality more than θ, and resubmit until acceptance. We formally define this central concept.

Definition 3.2 (De Facto Threshold). Consider a memoryless conference with acceptance rule ϕ and noiseless authors. A value θ such that a rational author submits a paper of quality $Q = q$ if $q > \theta$ and does not submit any papers with $q < \theta$ is called a *de facto threshold*.[11]

Notice that Proposition 3.1 shows that a de facto threshold always exists in our model.

3.2 Threshold Acceptance Policies and the Resubmission Gap

Threshold acceptance policies (see Definition 2.1) comprise a very natural class of policies for a conference to apply. Recall that they accept all papers whose posterior (based on the reviews) expected quality strictly clears some threshold, and reject all papers whose posterior expected

[11] This is because all such papers will be submitted until accepted, so de facto, the conference will accept all such papers. Note that for a paper with $q = \theta$, the author will decide whether to submit or not in favor of the conference.

quality falls short of the threshold. We show that every possible de facto threshold can be induced by a threshold acceptance policy.

Proposition 3.3. *Let θ be any de facto threshold. Then, there exists a $\hat{\tau}$ and a threshold acceptance policy with threshold $\hat{\tau}$ such that a rational noiseless author will submit if and only if $Q \geq \theta$.*

Due to Proposition 3.3, we can restrict our focus to acceptance *threshold* policies; this is because for any monotone policy, Proposition 3.1 implies the existence of a de facto threshold, and thus, Proposition 3.3 implies the existence of a threshold policy for which the author best-responds in the same way.

While Proposition 3.3 implies the existence of an acceptance threshold τ inducing the desired submission threshold θ, these two thresholds will typically be different. The threshold τ can be interpreted as the "declared" quality goal of the conference, attained in isolation. θ is the actual quality of the conference, taking into consideration resubmissions and reviewing noise. The difference is a key concept we study, and term the "resubmission gap."

Definition 3.4 (Resubmission Gap). Consider a memoryless conference with a threshold acceptance policy with threshold τ, and a noiseless author. Let θ be the author's submission threshold in response to the conference's policy. The difference between the acceptance threshold τ and the de facto threshold θ is called the *resubmission gap*.[12]

We also note that the optimal policy can be quite cleanly characterized.

Proposition 3.5. *The conference's optimal policy induces a de facto threshold of $\theta = 0$.*

Combining Proposition 3.5 with Proposition 3.1 allows us some insights into the right acceptance threshold for a conference aiming for maximum utility: under the chosen threshold, $P_{\text{acc}}(\tau, 0) = \rho$. This means that the more attractive the conference is (i.e., the smaller ρ), the more likely borderline papers must be rejected. Because borderline papers are resubmitted until accepted, this means that for the same optimal quality, an attractive conference (or patient authors) leads to higher reviewing load. We will investigate this phenomenon in significantly more depth in Section 4.3.

3.3 Continuous Model with Additive Noise

To derive some further insights, we now focus on the continuous model with additive noise and $m = 1$ review only. Importantly, recall that the noise distribution is independent of the true underlying quality, and has zero mean. That is, for any quality $Q = q$, the reviewer observes a signal of $q + X$ where the distribution of the review noise, $F^{(r)}(X)$, is independent of q. Under this model, the resubmission gap of a threshold acceptance policy does not depend on its threshold τ.

Proposition 3.6. *Given an acceptance threshold τ, in the continuous model with a single review and additive noise drawn from $F^{(r)}$, the de facto threshold is the unique solution to the equation $\rho = 1 - F^{(r)}(\tau - \theta)$. In particular, the resubmission gap $\tau - \theta$ is independent of τ and the prior distribution $\boldsymbol{p}$ over paper qualities.*

Substituting $\theta = 0$ from Proposition 3.5 into Proposition 3.6, we obtain the following.

Proposition 3.7. *For the continuous model with a single review and additive noise drawn from $F^{(r)}$ (independently of the true paper quality), the conference's optimal threshold τ^* is $\tau^* = (F^{(r)})^{-1}\left(\frac{V-1}{V-\eta}\right)$. Then, the author submits (and resubmits until accepted) the paper if and only if the quality is non-negative: $Q \geq 0$. The resulting (maximum) utility for the conference is $U^{(c)} = \int_0^\infty q\, dp(q)$.*

[12] The name reflects that there is a gap between the "intended" and "de facto" quality of accepted papers, caused by the fact that authors are free to resubmit papers.

First, note that Proposition 3.7 also implies that the quality-maximizing acceptance threshold does not depend on the distribution $\boldsymbol{p}$ of the papers' qualities. Second, notice that a large conference value V (high prestige) or a large discount factor η (patient authors) encourages authors to consistently resubmit bad papers. This leads to a large resubmission gap: the conference has to set a much higher bar to sufficiently discourage such resubmissions, and will reject many good papers repeatedly before they are finally accepted. In contrast, when the conference is not attractive enough or authors are not patient enough, the conference has to lower the acceptance threshold even below the de facto threshold, to provide strong assurance to good papers that they will be immediately accepted. However, this only occurs for $V < 2$; otherwise, the resubmission gap is non-negative for authors with any level of patience.

4 NOISELESS AUTHORS: TRADEOFFS AND ACCEPTANCE RATE

In this section, we build on the fundamental concepts of threshold policies, de facto thresholds, and resubmission gap, to undertake a more in-depth investigation of the tradeoffs a conference may face. In particular, we consider the tradeoff between conference quality and review burden on the community, and between acceptance thresholds and acceptance rate.

4.1 Continuous Models Studied

Here (and in the following sections containing discussions on the continuous model), we use the following special cases of our continuous model as examples for our analysis and plots.

The $(\sigma, \boldsymbol{p}, m, V, \eta)$-*Gaussian model* is a continuous model with noiseless authors; the noise for each review is drawn from a Gaussian distribution $F^{(r)} = \mathcal{N}(0, \sigma)$. The parameters $\boldsymbol{p}$, m, V, and η are, as before, the prior, the number of solicited reviews, the value of the prestigious conference, and the discount factor, respectively. The $(\sigma, \mu_{\boldsymbol{p}}, \sigma_{\boldsymbol{p}}, m, V, \eta)$-*Double Gaussian model* is a Gaussian model with the prior $\boldsymbol{p} = \mathcal{N}(\mu_{\boldsymbol{p}}, \sigma_{\boldsymbol{p}})$.

4.2 Tradeoff between Conference Quality and Review Burden: QB-tradeoff

We first consider the tradeoff between the conference quality and the review burden. We focus on the average number of reviews per *paper* (including papers that were never submitted, and thus incurred 0 reviews) as the relevant measure of the review burden.

Intuitively, if there are many borderline papers, whose quality is near the acceptance policy threshold τ, they may go through many rounds of resubmission and increase the review burden. Raising or lowering the threshold slightly might lead to a different borderline regime with a smaller fraction of papers on the border. We already saw that the conference quality is maximized by a de facto threshold of 0. Thus, changing the threshold will come at a cost to the conference, either by losing out on some good papers, or by accepting some bad papers. In the extremes, rejecting everything will lead to a review burden of 0, and accepting everything will lead to a review burden of 1. The *QB-Tradeoff* traces how the review burden and conference quality jointly vary across all possible acceptance policy thresholds.

We would like to understand the Pareto frontier of the QB-tradeoff over acceptance policies. That is, fixing all the parameters except the acceptance threshold, we want to understand which thresholds are Pareto optimal.

Notice that deviations from the optimal de facto threshold of 0 in *either* direction could be Pareto optimal. First, the conference can decrease the threshold to accept some negative-quality papers, in order to accept the positive-quality papers in fewer rounds; alternatively, the conference can increase the threshold to give up on some borderline papers with positive quality which might otherwise take a large number of rounds until acceptance. Clearly, which intervals of strategies are Pareto optimal depends on the distribution of paper quality. For example, if there is a substantially

larger number of borderline papers with negative quality than positive quality, marginally lowering the threshold will both degrade conference quality and increase review burden, but marginally increasing the threshold will decrease the review burden though still degrade conference quality. The latter will be Pareto optimal while the former will not.

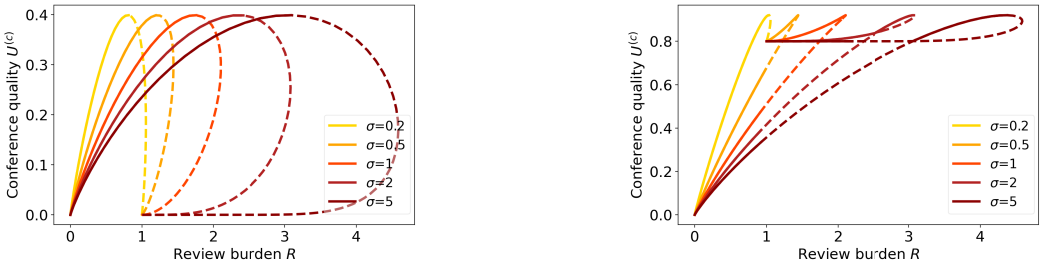


Fig. 1. (a) shows the QB-tradeoff of a $(\sigma, \mu_{\boldsymbol{p}} = 0, \sigma_{\boldsymbol{p}} = 1, m = 1, V = 5, \eta = .7)$-*Double Gaussian model*. (b) shows the QB-tradeoff of a $(\sigma, \mu_{\boldsymbol{p}} = .8, \sigma_{\boldsymbol{p}} = 1, m = 1, V = 5, \eta = .7)$-*Double Gaussian model*. In each case, the review quality σ is varied over five discrete options. For each σ, the curve shows the possible QB-tradeoffs as the acceptance threshold is varied continuously. The Pareto frontier is shown with solid lines, while dominated points are shown with dashed lines.

Fig. 1 maps the QB-tradeoff for various settings. In each setting, there is a point at (0, 0) that corresponds to rejecting all submissions. As the threshold is decreased, high-quality papers start being submitted, increasing both the conference quality and review burden. When the de facto threshold is 0, conference utility is maximized. Subsequently, a further decrease in the threshold leads to more low-quality papers being accepted, lowering the conference quality. We see here that the effect on the review burden is mixed: sometimes it increases while other times it decreases. As the threshold decreases further, in each case, we see a point with review burden 1 where all the papers are accepted. At that point, the conference quality is 0 in panel (a), but .8 in panel (b) due to the different priors on paper quality.

In Panel (a), the Pareto optimal de facto thresholds are always greater than 0. This is because, for any $\theta > 0$, the de facto thresholds of θ and $-\theta$ have the same conference quality, but θ will have a smaller review burden because the number of submitted papers overall is smaller, and the number of submitted papers just above the boundary will also be smaller.

Panel (b) contains some interesting observations. First, for $\sigma = 0.2$, 2 and 5, marginally increasing the de facto threshold from 0 remains Pareto optimal, while marginally decreasing it is Pareto dominated. However, the opposite is true for $\sigma = 0.1$ and 1. This is because, in all cases, when the threshold decreases, the number of submitted papers increases and the number of papers just above the borderline (that require more rounds of submission in expectation) decreases. However, the rate at which the latter decreases depends non-monotonically on the review quality. Second, notice that in the four largest settings of σ, some Pareto optimal thresholds lie both above and below the de facto threshold of 0. Thus, for all curves in panel (a) and $\sigma = 0.2$ in panel (b), we have that the Pareto optimal QB-tradeoffs either optimize the conference quality, or increase the threshold to trade off conference quality for a reduced review burden. However, for $\sigma = 0.5$ and 1 in panel (b), there exists a $\theta_0 > 0$, such that the Pareto optimal QB-tradeoffs either 1) only let in select papers with quality at least θ_0 or 2) keep out the bad papers (with quality less than some threshold $\theta \leq 0$). Here, having a de facto threshold of $0 < \theta < \theta_0$ is Pareto dominated by some threshold of $\theta' < 0$ which yields the same conference quality, but at a lower review burden. Finally,

for $\sigma = 2$ and 5 in panel (b), there are three ranges of Pareto optimal QB-tradeoffs! That is, there exists $\theta_2 < 0 < \theta_1 < \theta_0$ such that the Pareto optimal intervals either 1) only let in select papers for some threshold $\theta \geq \theta_0$; 2) keep out the really bad papers by setting the threshold to be $\theta \leq \theta_2$; or 3) maximize conference quality despite a relatively high review burden with $0 \leq \theta \leq \theta_1$.

4.3 Dominating QB-tradeoffs

In comparing the different QB-tradeoff curves of Figure 1, we observe that any curves corresponding to higher-quality (i.e., lower variance) reviews dominate similar curves corresponding to lower-quality reviews. This is not a coincidence. First, in general, we show that a curve is dominated if the corresponding conference value is higher or authors are more patient. Furthermore, we show that in the continuous model with a general family of noise distributions, a curve is dominated if the corresponding review quality is higher.

4.3.1 Conference Value and Agents' Discount Factor.

Proposition 4.1. *Consider two settings that are identical except that they have attractiveness factors ρ and $\rho' < \rho$, respectively. Then, the QB-tradeoff curve of the first setting dominates the QB-tradeoff curve of the second.*

Proposition 4.1 directly implies that either increasing the conference value V or increasing the authors' discount factor will harm the QB-tradeoff in the sense that it will be dominated by the original setting. Note that this result generally holds across the models that we considered and is also confirmed by our analysis based on real data in Appendix B. This observation has a range of practical implications which are discussed in Section 5.1. This observation has a range of practical implications which are discussed in Section 5.1.

4.3.2 Noise. Recall that in comparing the different curves of Figure 1, we observe that the higher the review quality (i.e., the lower the variance), the better the Pareto frontier. We next show that this is not a coincidence and holds not just for Gaussian noise in the reviews, but for any family of reviewer noise distributions in which the noise is scaled in a certain natural way. This notion is captured by the following definition:

Definition 4.2. Let X be a 0-mean unit-variance random variable with cdf F_X. We define the *scaling white noise family* with parameter $b > 0$ to have the cdf $F_b^{(r)}(x) = F_X\left(\frac{x}{b}\right)$, i.e., the cdf of the distribution of $b \cdot X$. As before, F_X is assumed to be invertible.

The following proposition implies that lower review noise among a scaling white noise family implies a higher acceptance rate for any paper above a fixed de facto threshold.

Proposition 4.3. *Let X be a 0-mean unit-variance random variable with cdf F_X. Consider two settings in the continuous model with memoryless threshold review policies that are identical except that the first has reviewer noise distributed according to $F_{b_1}^{(r)}(x)$ and the second according to $F_{b_2}^{(r)}(x)$ where $b_1 > b_2$. The QB-tradeoff curve in the first setting dominates the QB-tradeoff curve in the second.*

4.4 Acceptance Rate

One may suspect that the higher the threshold, the more selective the conference, so the lower the acceptance rate. But this is not always true. The reason is self selection by authors of weaker papers, who may not submit in the first place. As a result, those papers will not be rejected.

We first develop some mathematical tools to help us reason about the interaction between the selectivity of the conference (the de facto threshold) and the acceptance rate.

Let τ be the acceptance threshold, and θ the corresponding de facto threshold. As before, let $P_{\text{acc}}(\tau, q)$ be the probability that a paper of quality $Q = q$ is accepted at the conference. In round t, the total resubmission "density" of papers with quality $q \geq \theta$ is equal to $p(q) \cdot \sum_{j=0}^{t} (1 - P_{\text{acc}}(\tau, q))^{t-j}$. As t gets larger,[13] this converges to $p(q)/P_{\text{acc}}(\tau, q)$. Of these papers, a $P_{\text{acc}}(\tau, q)$ fraction will be accepted in each round. Hence, the acceptance rate converges to

$$\gamma = \frac{\text{Number of papers accepted this round}}{\text{Number of papers submitted this round}} = \frac{\int_{\theta}^{\infty} P_{\text{acc}}(\tau, q) \cdot p(q)/P_{\text{acc}}(\tau, q) dq}{\int_{\theta}^{\infty} p(q)/P_{\text{acc}}(\tau, q) dq}$$

$$= \frac{\int_{\theta}^{\infty} p(q) dq}{\int_{\theta}^{\infty} p(q)/P_{\text{acc}}(\tau, q) dq} = \frac{\int_{\theta}^{\infty} p(q) dq}{\int_{\theta}^{\infty} p(q)/(1 - F^{(r)}(\tau - q)) dq}. \tag{2}$$

We now use Eq. (2) to intuitively reason about how the prior distribution $\boldsymbol{p}$ affects the acceptance rate, γ. Notice that it is the papers with low acceptance probabilities which disproportionally decrease the acceptance rate. This is because they add only their mass to the numerator, but add their mass scaled by $1/P_{\text{acc}}(\tau, q)$ to the denominator. Thus, intuitively, if there is a z fraction of papers with quality at least θ that are "near" θ, and the other $1 - z$ fraction has a very high acceptance probability, the acceptance rate can be approximated by $\frac{1}{z/P_{\text{acc}}(\tau,\theta)+(1-z)}$. Notice that this is decreasing with z: the larger z, the smaller the acceptance rate.

The measurement of z intuitively resembles the *hazard rate* of a distribution, which is defined as $\frac{f(x)}{1-F(x)}$, where f is the probability density function, and F is the cumulative distribution function. Similar to z, the hazard rate measures the probability of a paper on a boundary at x, $f(x)$, relative to the mass of papers larger than x, $1 - F(x)$. The hazard rate is known to be monotone for thin-tailed distributions, like the Gaussian distribution. Conversely, the hazard rate is known to be non-monotone for heavy-tailed distributions, like the Cauchy distribution. Finally, the hazard rate is known to be (eventually) constant for the Laplace distribution.

Using the above intuition, we might expect the acceptance rates to decrease for the Gaussian prior. This is because Gaussian distributions have a monotone (increasing) hazard rate; thus, z is likely to increase and drive down the acceptance rate. Additionally, we might expect the acceptance rates to increase for the Cauchy distribution. This is because Cauchy distributions have a non-monotone hazard rate; thus, z is likely to decrease at some point and drive up the acceptance rate. Finally, we might expect that the acceptance rate is relatively constant after some point for the Laplace prior. Fig. 2 gives evidence to support this intuition. We find it remarkable that despite the heuristic reasoning "by analogy" to the hazard rate, we exactly observe these outcomes for the paradigmatic distributions in their respective classes.

Another interesting observation is that the quality distribution of submitted papers is not a good reflection of the prior quality distribution of papers, even conditional on being above the de facto threshold. The reason is that papers nearer the de facto threshold need to be submitted more times (on average) before being accepted than higher-quality papers. Therefore they are over-represented among the submitted papers.[14] However, by Eq. (2), the quality distribution of accepted papers is an accurate reflection of the prior quality distribution of papers conditional on being above the de facto threshold.

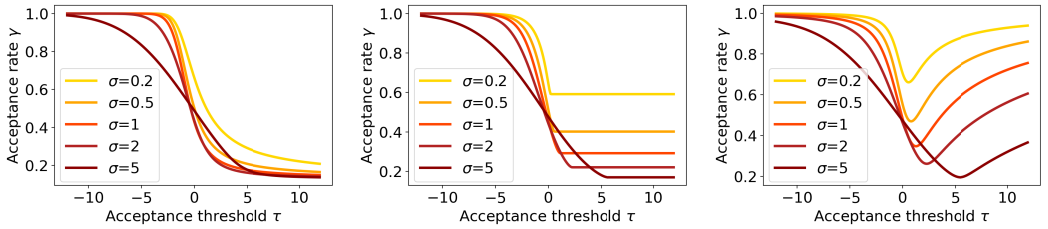
5 DISCUSSION, FUTURE WORK, AND CONCLUSION

5.1 Implications

Here, we suggest some possible interpretations and lessons that might be learned from our analysis.

[13] Taking t large enables us to take into account the gull history of previously rejected papers that are resubmitted.

[14] This aligns with many reviewers' observation in the real world that many of their assigned papers seem to be borderline.



(a) Gaussian prior. (b) Laplace prior. (c) Cauchy prior.

Fig. 2. The acceptance rate vs. acceptance threshold under different prior distributions. Here, the noise distribution $F^{(r)}$ is fixed as a normal distribution with zero mean and standard deviation of 1. Three types of prior distributions of the paper quality, all with zero mean and standard deviation (scaling factor) σ, are considered: (a) the Normal distribution; (b) the Laplace distribution; and (c) the Cauchy distribution.

Resubmission Gap. Perhaps the cleanest result concerns the resubmission gap. If a conference operates like our idealized model with noiseless authors, then we would observe the following: *Every paper ever submitted to the prestigious conference is eventually accepted; however, at any given conference, many papers will be rejected, and thus need to be submitted multiple times.* With many parameters, the vast majority of papers are rejected. On the surface, this sounds like a dystopian bureaucracy.

The frequently proposed and obvious reaction is to accept all acceptable papers the first time, without making them resubmit multiple times. This would allegedly decrease the reviewing load (because each paper would only be reviewed once), without affecting the quality of the conference because all of the papers were going to be accepted anyway. Superficially, this seems very reasonable.

However, our model warns that this is an unlikely outcome. Instead, by lowering its acceptance threshold, the conference would also lower its de facto threshold. While the papers being currently submitted could be overwhelmingly accepted in one round, it would invite more, lower-quality submissions. These lower-quality submissions would themselves be repeatedly submitted, and so the review burden would not necessarily decrease, and could even increase.

Of course, this is a not a perfect reflection of reality. In particular, because authors are not all aware of their papers' qualities, some submissions are of low quality and are very unlikely to ever be accepted. However, the present authors' experience is that for some prestigious conferences, the situation is similar to a large extent: the pool of submitted papers has been self-selected to those that will eventually appear in a good venue. Additional relevant modeling "blinds spots" are that the review noise for some papers may be different from others, agents may have different levels of patience and different values for their paper being accepted, and the distribution of paper qualities may react to the acceptance policy. However, it is not clear that any of these model limitations fundamentally challenge the insight that there is a gap between the quality of papers implied by the threshold acceptance policy and the types of papers submitted and eventually accepted.

Threshold Policy and QB-tradeoff. In the preceding discussion of the resubmission gap, it was not clear what happens to the review burden as the acceptance policy becomes more strict or lenient. This is because it depends on the prior distribution of paper qualities. This effect was studied in Section 4.2, where we identified three possible Pareto optimal acceptance policies: 1) accept only a few top papers; 2) accept all worthy papers; 3) accept nearly all papers, only attempting to weed out the worst. All of these policies can be easily identified in practice.

The advantage of policies 1) and 3) is they tend to require less reviewing. Policy 1) because few papers are submitted and policy 3) because few papers are submitted more than once. The advantage of policy 2) is that it maximizes the conference quality, though often at the expense of a high review burden.

In our analysis, however, the policy 3) is Pareto optimal only if the prior of paper qualities contains mostly positive papers. If, instead, the paper quality prior is a unimodal distribution centered at 0, then such a policy with a negative de facto threshold will never be Pareto optimal. We note that in the data learned from ICLR displayed in Fig. 4, all the acceptance policies which admit negative utility papers are Pareto dominated. (However, this is not always the case for the models we learned from ICLR data for other years or parameters).

Conference Value and Discount Factor. We showed, in the noiseless author setting, that increasing a conference's prestige and thus value, or increasing the discount factor, creates a strictly worse QB-tradeoff curve. This has a range of practical implications.

First, having very high prestige conferences creates a higher reviewing burden. In our model, this happens regardless of the acceptance policy; however, more realistically, lowering the acceptance policy causally decreases the value of the conference which our model does not account for.

Second, policies that burden the author may improve the QB-tradeoff. Examples include long review times, rebuttal periods, or onerous formatting requirements. This will intuitively allow the conference to decrease their acceptance threshold while keeping the same de facto threshold, and thus may decrease the review load without impacting conference quality. Essentially, such policies artificially internalize the negative externality of imposing reviews upon others.

Conversely, our model predicts that well-meaning efforts to reduce these burdens may very well worsen the QB-tradeoff. Proposed and executed reforms include: decreasing the required time to review papers for a given conference and a "desk reject" phase, where papers that appear subpar are quickly returned to authors without review. These policies artificially increase the time discount η by decreasing the time between submissions. This forces the conference to increase its acceptance policy threshold if it would like to maintain the same de facto threshold (and thus maintain conference quality).

Of course, it should be emphasized that these observations are not to be taken as recommendations: in particular, as we discuss in Section 5.2, our analysis essentially ignores the authors' utilities, and a longer time until acceptance or more burden to submit leads to a decrease of this utility. The tradeoff should naturally include all three concerns, and our observations should be interpreted as emphasizing one aspect that may not have been considered enough in the past.

Acceptance Rate. We show that empirically whether the acceptance rate increases, decreases, or remains steady as the acceptance threshold increases is a function of the hazard rate of the prior paper quality distribution. This warns against using the acceptance rate as a signal of quality. At issue is that for certain priors of paper quality, as the selectivity increases, the fraction of papers near the boundary decreases.

Additional factors may complicate this picture. For example, higher-quality papers may have a different review noise distributions than lower-quality papers.

We also noted that in our noiseless model, the quality distribution of accepted papers equals the prior paper quality distribution conditional on being greater than the de facto threshold. However, we note that our ICLR data set was instead learned from the submitted papers, so it may overcount borderline papers while under-counting low-quality papers.

Quality vs. Quantity of Reviews. Our results here discourage the strategy of soliciting a large number of reviews per paper. In particular, any number of solicited reviews larger than 3 greatly

burdens the review system but is unlikely to bring enough benefits to the conference quality. Instead, our model predicts that a small number of solicited reviews, even with one review per paper, can be optimal if the conference is able to find the optimal acceptance threshold. The intuition is that when authors know the quality of their papers well enough, any number of solicited reviews, as long as it is combined with the optimal acceptance threshold, can take the advantage of authors' self-selection such that only the desired papers are submitted, and eventually accepted.

A larger number of reviews may nonetheless be desirable due to several real-world considerations: first, more reviews may decrease the average number of times a paper needs to be resubmitted, helping authors; second, our reviewer error model does not capture the situation where there may be a "fatal flaw" that only an astute reviewer observes; and finally, more reviewers can provide more feedback. Relatedly, in Fig. 9, we observe that soliciting only one review per paper may slow the progress of authors learning about the quality of their papers based on historical reviews; this results in low conference quality. Moreover, our model assumes that reviews are i.i.d., which is not true in reality when the reviewers can communicate (after the rebuttal). The integrated review signals after communication may become much more informative than aggregating each of them as an i.i.d. review; this may significantly benefit the strategy of soliciting a large number of reviews.

Institutional Memory. Our results indicate that the idea of having historical reviews follow submissions, or more generally, allowing the conference to limit the number of times a paper can be submitted, can help reduce the review burden but the effect is only marginal. The reason is partially that even a memoryless acceptance policy, as long as it is optimally designed, can achieve a reasonably good tradeoff between the conference quality and the review burden.

We note that due to computational concerns, our conclusion here is largely based on the ABM experiments in the simplest binary model. Given that there seems to be general resistance to the idea of review-following (perhaps for fear that a bad, but erroneous, review may bring bias to the following reviews and doom a paper's chances for a long time), even if these modest gains generalize, this idea is unlikely to be a game changer.

5.2 Limitations and Future Directions

Author Utility. A big limitation of our work is that we do not explicitly measure author utility. Although our model of author utility makes sense from the point of view of modeling the authors' local *decisions*, it does not capture the authors' actual utility in a holistic sense. For example, if the conference accepted everything, then every author's utility would be V according to the model (and indeed, authors would prefer having their papers accepted); however, the prestige of a conference is also derived from its selectivity, so the authors of high-quality papers would not be happy with this outcome. In reality, V is a long-term function of the quality of papers appearing in that venue, a fact that we do not include in the model.

We do note that, typically in our analysis, for a fixed de facto threshold, the average number of submissions for a paper is closely related to the review burden and is a reasonable proxy for the author utility. Thus, the authors' interests are, to a limited extent, present. However, this relationship breaks down on occasion, for example, when discussing the number of reviews that should be solicited per submission. Here, the conference's cost is measured in reviews, while the author's cost is measured in rounds of submission. Because m is not fixed, these are not the same.

Future work could make author utility a first-order concern, in part by modeling conference value as dependent on the quality of the papers at the conference.

A Single Prestigious Conference. Another limitation of this work is that we assume a single prestigious conference. As mentioned, this can model several prestigious conferences that are more or less cooperating to uphold community standards. In reality, there is an ecosystem of conferences

and not all of them are either top-tier or sure bets. In such a setting, our analysis could model the decision to submit to a top-tier or second-tier conference. The utility for submitting to the second tier could be normalized to 1. The issue with this, however, is that the second-tier conference still needs to review its submissions. Other analyses (see related work in Section 1.2) have considered venues of different values.

Furthermore, our model omits competition between conferences. For example, conferences may compete to attract more papers by attempting to increase their quality, making their acceptance policy more predictable, creating a faster turnaround time, etc.

Heterogeneity. We also did not model various sources of heterogeneity in the process. For example, the qualities of reviews are not uniform, and different authors have different levels of patience (e.g., a Postdoc who will be on the job market vs. a first-year student or a tenured faculty member). We are also not modeling the effects of biases that may impact different researchers disproportionately. The uneven impact on different author populations would be made more complex by co-authorship. As with the previously mentioned endogenous review quality, a main difficulty would be that this model would have more parameters, and as such require learning/setting them, which could lead to arbitrary choices.

Additional Feedback Loops. There are several feedback loops that we disregard. We model paper quality as exogenous; however, in reality, it is largely determined by authors who decide how much effort or time to expend improving their papers. Authors often write a paper to target particular venues. As the venues change their policies, the underlying distribution of paper qualities is likely to change. Moreover, authors may improve their papers in response to reviewer feedback. Additionally, as mentioned already, in our model the acceptance policy does not impact the conference value, and the review burden does not impact the quality of the reviews or the amount of time authors (who are typically also the reviewers) spend writing papers. As mentioned in Section 1.2, several results do try to model these complexities and provide insights with agent-based simulations [5, 17, 26].

Here, we focus on a simpler and cleaner model than can be afforded when including these complexities. Apart from the analysis being more difficult, it is often difficult to know precisely how these feedback loops function, which may lead to even greater uncertainty of the accuracy of a model, and any insights derived from it.

Modeling Paper Quality. Finally, it should be noted that papers do not have an "objective" quality that can be projected to a single dimension or even multiple dimensions. One approach, which adds minimal complexity along these lines, is to distinguish between different types of poor-quality papers. Some papers may be deemed low quality because they are methodologically flawed; others because their contributions may be incremental. The first of these might be easier for reviewers to detect and agree upon than the second.

5.3 Conclusion

Despite the fact that there may be no agreed-upon objective metric, still the rigor of peer review is important to a healthy future of academic research. We avoid concrete recommendations, as many quantities are highly depend on model parameters, and on unmodeled real-life effects. We hope that our theory (and simulations) can steer the discussion, uncover parameters to focus on, and inform decision makers. We believe that the focus on resubmission gap, and the importance of reviewing quality over quantity, are important points to start a discussion in the community which may not have been as easily identified without studying a model like ours.

REFERENCES

[1] 2020. ICLR 2020 review data. https://github.com/shaohua0116/ICLR2020-OpenReviewData.

[2] 2021. ICLR 2021 review data. https://github.com/evanzd/ICLR2021-OpenReviewData.

[3] Stefano Allesina. 2012. Modeling peer review: An agent-based approach. *Ideas in Ecology and Evolution* 5, 2 (2012).

[4] Ammar Ammar and Devavrat Shah. 2012. Efficient rank aggregation using partial data. *ACM SIGMETRICS Performance Evaluation Review* 40, 1 (2012), 355–366.

[5] Federico Bianchi, Francisco Grimaldo, Giangiacomo Bravo, and Flaminio Squazzoni. 2018. The peer review game: an agent-based model of scientists facing resource constraints and institutional pressures. *Scientometrics* 116, 3 (2018), 1401–1420.

[6] Domenic V. Cicchetti. 1991. The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation. *Behavioral and brain sciences* 14, 1 (1991), 119–135.

[7] Stephen Cole, Jonathan R. Cole, and Gary A. Simon. 1981. Chance and consensus in peer review. *Science* 214, 4523 (1981), 881–886.

[8] Wade D. Cook, Boaz Golany, Michal Penn, and Tal Raviv. 2007. Creating a consensus ranking of proposals from reviewers' partial ordinal rankings. *Computers & Operations Research* 34, 4 (2007), 954–965.

[9] Alexander Philip Dawid and Allan M. Skene. 1979. Maximum likelihood estimation of observer error-rates using the EM algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 28, 1 (1979), 20–28.

[10] Rafael D'Andrea and James P. O'Dwyer. 2017. Can editors save peer review from peer reviewers? *PloS One* 12, 10 (2017), e0186111.

[11] Robert L. Ebel. 1951. Estimation of the reliability of ratings. *Psychometrika* 16, 4 (1951), 407–424.

[12] Thomas Feliciani, Junwen Luo, Lai Ma, Pablo Lucas, Flaminio Squazzoni, Ana Marušić, and Kalpana Shankar. 2019. A scoping review of simulation models of peer review. *Scientometrics* 121, 1 (2019), 555–594.

[13] David Gill and Daniel Sgroi. 2012. The optimal choice of pre-launch reviewer. *Journal of Economic Theory* 147, 3 (2012), 1247–1260.

[14] Daniel M. Herron. 2012. Is expert peer review obsolete? A model suggests that post-publication reader review may exceed the accuracy of traditional peer review. *Surgical Endoscopy* 26, 8 (2012), 2275–2280.

[15] Steven Jecmen, Hanrui Zhang, Ryan Liu, Nihar B. Shah, Vincent Conitzer, and Fei Fang. 2020. Mitigating Manipulation in Peer Review via Randomized Reviewer Assignments. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin (Eds.). https://proceedings.neurips.cc/paper/2020/hash/93fb39474c51b8a82a68413e2a5ae17a-Abstract.html

[16] Sampath Kannan, Mingzi Niu, Aaron Roth, and Rakesh Vohra. 2021. Best vs. All: Equity and Accuracy of Standardized Test Score Reporting. (2021). arXiv preprint arXiv:2102.07809.

[17] Michail Kovanis, Raphaël Porcher, Philippe Ravaud, and Ludovic Trinquart. 2016. Complex systems approach to scientific publication and peer-review system: development of an agent-based model calibrated with empirical journal data. *Scientometrics* 106, 2 (2016), 695–715.

[18] Michail Kovanis, Ludovic Trinquart, Philippe Ravaud, and Raphaël Porcher. 2017. Evaluating alternative systems of peer review: a large-scale agent-based modelling approach to scientific publication. *Scientometrics* 113, 1 (2017), 651–671.

[19] Josh Lerner and Jean Tirole. 2006. A model of forum shopping. *American economic review* 96, 4 (2006), 1091–1113.

[20] Bryan D. Neff and Julian D. Olden. 2006. Is peer review a game of chance? *BioScience* 56, 4 (2006), 333–340.

[21] Ritesh Noothigattu, Nihar B. Shah, and Ariel Procaccia. 2018. Choosing how to choose papers. (2018). arXiv preprint arxiv:1808.09057.

[22] Paul J. Roebber and David M. Schultz. 2011. Peer Review, Program Officers and Science Funding. *PLoS ONE* 6, 4 (2011), e18680.

[23] Anna Rogers and Isabelle Augenstein. 2020. What Can We Do to Improve Peer Review in NLP?. In *Findings of the Association for Computational Linguistics: EMNLP 2020, Online Event, 16-20 November 2020 (Findings of ACL)*, Trevor Cohn, Yulan He, and Yang Liu (Eds.), Vol. EMNLP 2020. Association for Computational Linguistics, 1256–1262. https://doi.org/10.18653/v1/2020.findings-emnlp.112

[24] Nihar B. Shah, Behzad Tabibian, Krikamol Muandet, Isabelle Guyon, and Ulrike Von Luxburg. 2018. Design and analysis of the NIPS 2016 review process. *Journal of Machine Learning Research* 19 (2018), 1–49.

[25] Lones Smith and Andrea Wilson. 2021. Accept this Paper. (2021). Manuscript in preparation. Slide deck available at https://econ.la.psu.edu/events/seminar-documents/accept2021.pdf.

[26] Flaminio Squazzoni and Claudio Gandelli. 2012. Saint Matthew strikes again: An agent-based model of peer review and the scientific community structure. *Journal of Informetrics* 6, 2 (2012), 265–275.

[27] Siddarth Srinivasan and Jamie Morgenstern. 2021. Auctions and Prediction Markets for Scientific Peer Review. (2021). arXiv preprint arXiv:2109.00923.

[28] Weijie J. Su. 2021. You Are the Best Reviewer of Your Own Papers: An Owner-Assisted Scoring Mechanism. arXiv:cs.LG/2110.14802

[29] Stefan Thurner and Rudolf Hanel. 2011. Peer-review in a world with rational scientists: Toward selection of the average. *The European Physical Journal B* 84, 4 (2011), 707–711.

[30] David Tran, Alex Valtchanov, Keshav Ganapathy, Raymond Feng, Eric Slud, Micah Goldblum, and Tom Goldstein. 2020. An Open Review of OpenReview: A Critical Analysis of the Machine Learning Conference Review Process. (2020). arXiv preprint arXiv:2010.05137.

[31] Jingyan Wang and Nihar B. Shah. 2019. Your 2 is My 1, Your 3 is My 9: Handling Arbitrary Miscalibrations in Ratings. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '19, Montreal, QC, Canada, May 13-17, 2019*, Edith Elkind, Manuela Veloso, Noa Agmon, and Matthew E. Taylor (Eds.). International Foundation for Autonomous Agents and Multiagent Systems, 864–872. http://dl.acm.org/citation.cfm?id=3331778

[32] Hanrui Zhang, Yu Cheng, and Vincent Conitzer. 2021. Classification with Few Tests through Self-Selection. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*. AAAI Press, 5805–5812. https://ojs.aaai.org/index.php/AAAI/article/view/16727

A PROOFS

A.1 Proof of Proposition 2.2

Proof of Proposition 2.2. Let $q' > q$. Consider the draws of s_i from $\boldsymbol{\beta}_q$ and s_i' from $\boldsymbol{\beta}_{q'}$. Because $\boldsymbol{\beta}_{q'}$ (strictly) first-order stochastically dominates $\boldsymbol{\beta}_q$, these draws can be coupled such that $s_i' \geq s_i$ always holds, and the inequality is strict with positive probability. Then, the inequality on acceptance probability directly follows from the definition of monotonicity of ϕ. □

A.2 Proof of Proposition 3.1

Proof of Proposition 3.1. The probability of acceptance for a paper of quality $Q = q$ is $P_{\text{acc}}(\phi, q)$, which is increasing in q by Proposition 2.2. The author submits the paper if and only if the expected utility of submitting is at least 1 (the value of the "sure bet" option). Since the author will face the same tradeoff in future rounds[15], she will make the same decision, so she will submit until acceptance. The expected utility can be obtained by the time-discounted sum over time periods of acceptance:

$$1 \leq \sum_{t \geq 1} V \cdot \eta^{t-1} P_{\text{acc}}(\phi, q) \cdot (1 - P_{\text{acc}}(\phi, q))^{t-1} = \frac{V \cdot P_{\text{acc}}(\phi, q)}{1 - \eta \cdot (1 - P_{\text{acc}}(\phi, q))}. \tag{3}$$

Solving this inequality for q, the author submits the paper if and only if $P_{\text{acc}}(\phi, q) \geq \rho$. By monotonicity of $P_{\text{acc}}(\phi, \cdot)$, the author submits if and only if $q \geq \theta$. □

A.3 Proof of Proposition 3.3

Proof. Consider the threshold policy $\phi_{\tau,0}$ with threshold τ which accepts a paper if and only if $\mathbb{E}[Q \mid \boldsymbol{s}] > \tau$. Let $f(\tau) := P_{\text{acc}}(\phi_\tau, \theta)$ be the probability that a paper of quality θ is accepted under the policy ϕ_τ.

Then, $\lim_{\tau \to -\infty} f(\tau) = 1 > \rho > 0 = \lim_{\tau \to \infty} f(\tau)$. Furthermore, $f(\tau)$ is a non-increasing function of τ, because the set of signal vectors $\boldsymbol{s}$ under which the conditional expected quality exceeds τ is monotone (with respect to inclusion). Therefore, there must exist a $\hat{\tau}$ such that $\lim_{\tau \to \hat{\tau}\uparrow} f(\tau) \geq \rho \geq \lim_{\tau \to \hat{\tau}\downarrow} f(\tau)$. Fix this $\hat{\tau}$.

If f is continuous at $\hat{\tau}$, then the threshold policy $\phi_{\hat{\tau},0}$ has $P_{\text{acc}}(\phi_{\hat{\tau},0}, \theta) = \rho$ by definition. Otherwise, let $z = \lim_{\tau \to \hat{\tau}\uparrow} f(\tau) - \lim_{\tau \to \hat{\tau}\downarrow} f(\tau) > 0$. We can infer that there must be a discrete probability of z for the event that $\mathbb{E}[Q \mid \boldsymbol{s}] = \hat{\tau}$, i.e., that $\text{Prob}_{\boldsymbol{s} \sim \boldsymbol{\beta}_\theta}[\mathbb{E}[Q \mid \boldsymbol{s}] = \hat{\tau}] = z$. We then consider the threshold policy $\phi_{\hat{\tau},r}$ with threshold $\hat{\tau}$ which conditioned on $\mathbb{E}[Q \mid \boldsymbol{s}] = \tau$ accepts a paper with probability $r := \frac{\rho - \lim_{\tau \to \hat{\tau}\downarrow} f(\tau)}{z}$. The overall acceptance probability of $\phi_{\hat{\tau},r}$ for a paper with quality θ is therefore

$$\begin{aligned} &\text{Prob}_{\boldsymbol{s} \sim \boldsymbol{\beta}_\theta}[\mathbb{E}[Q \mid \boldsymbol{s}] > \hat{\tau}] + \text{Prob}_{\boldsymbol{s} \sim \boldsymbol{\beta}_\theta}[\mathbb{E}[Q \mid \boldsymbol{s}] = \hat{\tau}] \cdot \frac{\rho - \lim_{\tau \to \hat{\tau}\downarrow} f(\tau)}{z} \\ &= \lim_{\tau \to \hat{\tau}\downarrow} f(\tau) + z \cdot \frac{\rho - \lim_{\tau \to \hat{\tau}\downarrow} f(\tau)}{z} = \rho. \end{aligned}$$

Thus, θ is the de facto threshold of the threshold policy $\phi_{\hat{\tau},r}$. □

A.4 Proof of Proposition 3.5

Proof of Proposition 3.5. Given a de facto threshold θ, the conference's quality is $\int_\theta^\infty q \, dp(q)$. This expression is maximized when $\theta = 0$, proving the optimality of this de facto threshold. □

[15] Crucially, a noiseless author does not learn any new information from rejection in previous rounds.

A.5 Proof of Proposition 3.7

Proof. Consider an acceptance threshold τ, and corresponding policy ϕ. Because we are considering a continuous model, the probability that the conditional expected quality of a paper is exactly τ is 0, so τ uniquely defines ϕ. By Proposition 3.1, the de facto threshold satisfies $P_{\text{acc}}(\phi, \theta) = \rho$.

In the continuous model with a single review, the conference will accept a paper with quality q if and only if the review s satisfies $s = q + x > \tau$. This happens with probability $P_{\text{acc}}(\phi, q) = 1 - F^{(r)}(\tau - q)$. Thus, θ solves $\rho = 1 - F^{(r)}(\tau - \theta)$.

A solution for θ exists because $\rho = \frac{1-\eta}{V-\eta} \in (0, 1)$ (because $\eta \in (0, 1)$ and $V > 1$) and $F^{(r)}(\cdot)$ is a distribution. Uniqueness of θ follows because we assumed $F^{(r)}(\cdot)$ to be invertible. □

A.6 Proof of Proposition 4.1

We first show the following lemma.

Lemma A.1. *Consider two settings with identical de facto thresholds, and thus the same conference quality. Let C and C' be the corresponding QB-tradeoff curves. If for every de facto threshold θ that neither accepts all papers nor rejects all papers, the review burden is strictly greater in the C' setting than in the C setting, then the QB-tradeoff curve C dominates the QB-tradeoff curve C'.*

Proof. Any point on the QB-tradeoff curve C' has a corresponding de facto threshold θ. Unless it is the de facto threshold that either accepts all papers or rejects all papers, it is Pareto dominated by the point on the QB-tradeoff curve C corresponding to the same de facto threshold because this point has, by assumption, strictly lower review burden and the same conference quality. □

The lemma can be straightforwardly applied to prove Proposition 4.1

Proof of Proposition 4.1. We will show that for every de facto threshold θ that neither rejects all papers nor accepts all papers, the review burden is strictly less in the first setting than in the second and then apply Lemma A.1.

Fix any de facto threshold $\theta \notin \{-\infty, +\infty\}$. By Proposition 3.3, both settings have a threshold acceptance policy with some thresholds τ and τ'. From Proposition 3.1, we know that ρ and ρ' are the probabilities with which a paper of quality θ is accepted in each setting. Thus, either 1) $\tau < \tau'$ or 2) we are in a knife-edge case where $\tau = \tau'$, $\text{Prob}_{\boldsymbol{s}}[U(\boldsymbol{s}) = \tau = \tau'] > 0$, and when the reviewer signal is such that $U(\boldsymbol{s}) = \tau$, the first setting rejects with strictly smaller probability than the second setting. Notice that in either case, every submitted paper is less likely to be accepted in the second setting than the first, and some strictly less. Thus, the expected number of times some papers are submitted will increase. However, the same papers are submitted in each case (those with quality exceeding the de facto threshold). Thus, the review burden is strictly greater in the second case than the first. □

A.7 Proof of Proposition 4.3

Proof. We will show that for every de facto threshold $\theta \notin \{-\infty, +\infty\}$, when the noise is distributed according to $F_b^{(r)}(x)$, the acceptance probability of a paper of quality $q \geq \theta$ is strictly increasing in b. That is, the smaller the reviewer noise, the more likely the paper is to be accepted. Then the result follows by applying Lemma A.1.

First, we relate the conference's acceptance threshold with the de facto threshold. Recall from Proposition 3.7 that the thresholds satisfy $F_b^{(r)}(\tau - \theta) = F_X\left(\frac{\tau-\theta}{b}\right) = \frac{V-1}{V-\eta}$.

Writing $C = F_X^{-1}\left(\frac{V-1}{V-\eta}\right)$ (a constant which does not depend on the scaling parameter b), this equation solves to $\tau = \theta + b \cdot C$.

For an acceptance threshold $\tau = \theta + b \cdot C$, a paper of quality $Q = q$ is accepted if and only if the review noise $b \cdot X$ is greater than $\tau - q$. However, this is equivalent to $X \geq \frac{\tau - q}{b} = \frac{\theta + b \cdot C - q}{b} = C - \frac{q - \theta}{b}$. Because $q - \theta$ is positive, the-right hand side strictly increases as b grows. Thus, X being greater than this value is strictly less likely as b grows. This proves the proposition. □

B NOISELESS AUTHORS AND REAL-WORLD PARAMETERS

In Section 3, we defined the key notions of de facto thresholds and resubmission gaps, and characterized them generically in terms of model parameters. In Section 4, we built on this foundation to obtain theoretical insights into Pareto-optimal QB-tradeoffs and the relationship between the acceptance threshold and acceptance rate. In this section, our goal is to investigate these basic phenomena on more realistic categorical data. We will be interested in a more theoretical model with binary qualities and signals (which allows for very clean interpretation of model parameters), and a realistic model with parameters learned from publicly available ICLR data.

Though our results in this section are derived analytically, we use simulations and figures to visualize them and omit cumbersome analytical formulations.

B.1 Discrete Models Studied

We define the following two special cases of our general model. Both have discrete sets of qualities $\boldsymbol{Q}$ and signals Σ. Both models are defined more generally, allowing for noisy authors, for use in later sections.

The $(\alpha, \beta, m, V, \eta)$-*binary model* is a categorical model in which there are two paper qualities $\{-1, +1\}$ and two review signals. One paper quality is referred to as negative ("bad papers"), and the other as positive ("good papers"). The prior is such that each paper is equally likely to be bad or good. Authors receive the correct signal about their papers with probability α and otherwise receive the opposite signal. Similarly, each reviewer receives the correct signal about his assigned paper with probability β and otherwise receives the opposite signal. The parameters m, V, and η are, as in general, the number of solicited reviews, the value of the prestigious conference, and the discount factor, respectively.

The $(\lambda_A, \lambda_R, m, V, \eta) - ICLR^{y,L}$ *model* is a categorical model learned from data. Specifically, the prior of paper quality $\boldsymbol{p}^*$ and the review signal distribution $\boldsymbol{\beta}^*$ are learned from the ICLR 2020 and 2021 open review datasets [1, 2]; for each dataset, a model is learned with paper quality sets of sizes $L = |\boldsymbol{Q}| \in \{4, 5\}$. The year of the dataset is denoted by y. Details of the method used for learning are described in Appendix E, where the learned parameters are shown in Table 1.

When varying the signal quality in models based on ICLR data, we use $\lambda_A \in [0, 1]$ for the author's signal and $\lambda_R \in [0, 1]$ for the reviewer's signal; the parameters control the probability with which the signal is drawn from the learned parameters $\boldsymbol{\beta}^*$ vs. a uniform distribution. $\lambda_A = 1$ ($\lambda_R = 1$) implies that the signal is drawn from $\boldsymbol{\beta}^*$ (the same for both authors and reviewers) while $\lambda_A = 0$ ($\lambda_R = 0$) implies that the signal is uniformly random. The parameters m, V, and η are again the number of solicited reviews, the value of the prestigious conference, and the discount factor, respectively. When we model noiseless authors, we simply remove the entry λ_A from the tuple and call this the $(\lambda_R, m, V, \eta) - \text{ICLR}^{y,L}$ noiseless-author model.

We remark that the distribution $\boldsymbol{\beta}$ of reviewer signals is informative in all of our inferred data sets, which we tested using an exhaustive comparison.

B.2 The Resubmission Gap in Categorical Models

We begin by juxtaposing the acceptance threshold and de facto threshold in a model whose parameters are learned from ICLR data, and in which each paper is reviewed $m = 3$ times. We study

this juxtaposition to understand the extent of the resubmission gap for a situation with realistic parameters for a real-world conference.

By Proposition 3.3, each de facto threshold can be achieved by a threshold policy with some acceptance threshold; we therefore focus on threshold acceptance policies. However, given that the models we consider here are categorical, a threshold τ may not uniquely determine a policy, namely, when there is a combination of reviewer signals which occurs with positive probability and induces posterior expected quality exactly τ. In order to associate a unique policy ϕ_τ with each $\tau \in \mathbb{R}$, we use the following convention: Let $\tau' < \tau''$ be two expected posterior qualities that arise with positive probability conditioned on signals, which are "adjacent" in the sense that no $\tilde{\tau} \in (\tau', \tau'')$ arises with positive probability as expected posterior quality. Then, we interpret a threshold of $\tau \in [\tau', \tau'']$ as the policy which accepts all papers of expected posterior quality larger than τ'', rejects all papers of expected posterior quality less than τ', and accepts papers of expected posterior quality exactly τ' with probability $\frac{\tau-\tau'}{\tau''-\tau'}$.

Recall from Proposition 3.5 that optimal policies for the conference are exactly those that induce a de facto threshold of $\theta = 0$ in the authors; that is, they induce authors to submit all papers of positive quality but no papers of negative quality. If there are multiple threshold acceptance policies all achieving the same de facto threshold of 0, we have the conference choose the most lenient of these, i.e., maximizing the acceptance probability. This is because the same set of papers is submitted and accepted eventually, but each is reviewed the smallest number of times.

Recall that by Proposition 3.1, a paper with quality $Q = q$ is submitted if $P_{\text{acc}}(\tau, q) > \rho = \frac{1-\eta}{V-\eta}$ and not submitted if $P_{\text{acc}}(\tau, q) < \rho$. Fig. 3 plots $P_{\text{acc}}(\tau, q)$ as a function of the acceptance threshold τ, for all $q \in Q$ in the categorical model. In this figure, there are four discrete paper qualities, two of which are negative. The optimal de facto threshold is just above the largest negative quality, which is $\hat{\theta} = -0.41$. The optimal acceptance threshold is determined by the intersection between the horizontal line $y = \rho$ and the curve $P_{\text{acc}}(\tau, \hat{\theta})$ (the orange curve), which implies that the optimum threshold τ^* is the lowest threshold to guarantee the maximum quality. By definition, the resubmission gap is then $\tau^* - \hat{\theta}$.

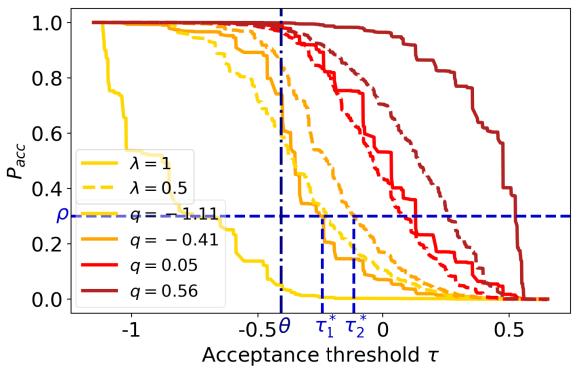
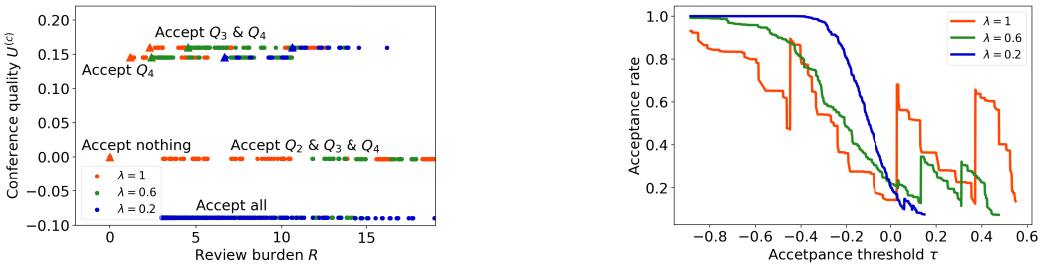


Fig. 3. The probability of acceptance $P_{\text{acc}}(\tau, q)$ as a function of the acceptance threshold τ in the $(\lambda_R, m = 3, V, \eta) - ICLR^{2020,4}$ *noiseless-author model*. In this example, $\rho = 0.3$, the optimal de facto threshold is $\theta = -0.41$. We consider two different levels of review quality, characterized by $\lambda_R = 1$ (higher quality) and $\lambda_R = 0.5$; the corresponding curves are shown in solid and dashed lines, respectively. The corresponding optimal acceptance thresholds are $\tau_1^* = -0.24$ and $\tau_2^* = -0.11$.

From Fig. 3, it is straightforward that increasing ρ and λ leads to a decrease in the resubmission gap. Recall that $\rho = \frac{1-\eta}{V-\eta}$. Hence, the resubmission gap is increasing in V and η, and decreasing in the review quality. Both observations are in line with our theoretical results in Section 3. Note that the curves in Fig. 3 are serrated because the expected quality of the conference only takes on a discrete set of values, and P_{acc} jumps when the de facto threshold passes one of these discrete values.

B.3 The Quality-Burden Tradeoff and Acceptance Threshold



(a) Quality-burden tradeoff

(b) Acceptance rate.

Fig. 4. In the $(\lambda_R, m = 3, V = 5, \eta = 0.7)$ – *ICLR*^{2020,4} *noiseless-author model*, (a) shows the different quality-burden tradeoffs; (b) shows the acceptance rate as a function of the acceptance threshold. For (a), the Pareto optimal points are marked with triangles for different review qualities. Note that the de facto threshold which accepts Q_2, Q_3 and Q_4 has expected conference quality -0.002, and thus is dominated by "accepting nothing." Similarly, "accept all" is dominated by "accept nothing."

As in the continuous model, we next want to understand the tradeoff between the conference's quality and the review burden on the community, as well as the relationship between the conference's acceptance rate, the review quality, and the acceptance threshold in the categorical model.

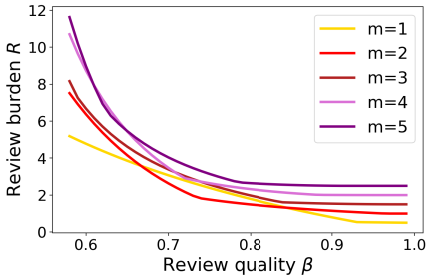
Fig. 4(a) shows the QB-tradeoff. Similar to the continuous case, the conference can reduce the number of reviews while giving up some quality, by increasing the threshold and only accepting the papers of highest quality. Different from the continuous model, the Pareto frontier of the categorical model is not a curve but a set of discrete points; this is due to the discrete nature of the quality space. Furthermore, in line with Proposition 4.3 (which, however, only applies to the continuous case), the Pareto optimal points improve with the review quality (as captured by λ), i.e., higher review quality allows for a lower review burden while achieving the same conference quality.

The acceptance rate, as shown in Fig. 4(a), has a sawtooth like shape as a function of the acceptance threshold. It jumps up at discrete points whenever the acceptance threshold reaches a level where the lowest remaining tier of papers will not be submitted any more. Subsequently, as the acceptance threshold increases further, the acceptance rate decreases, as the same papers are submitted, but more papers are rejected due to the higher threshold. Furthermore, keeping the acceptance threshold constant, higher review quality leads to *lower* acceptance rate when the acceptance threshold is relatively low, but to *higher* acceptance rate when the threshold is high. The intuition is that for low-quality reviews, there is some minimum $U(\boldsymbol{s})$, and once the threshold is below this, everything is immediately accepted and thus also submitted. However, the high-quality reviews have a lower minimum $U(\boldsymbol{s})$, and so for low thresholds, they will reject papers with poor reviews; but, nonetheless, all papers are submitted. When the acceptance threshold is

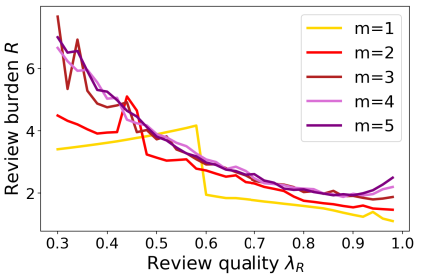
high, high-quality reviews lead to more careful self-selection than low-quality reviews, and thus lead to a higher acceptance rate.

B.4 The Optimal Number of Solicited Reviews

We next study the optimal number of reviews the conference should solicit in each round. Notice that an increase in the number of reviews per round may be counter-balanced by a decrease in the number of rounds until a paper is accepted. More specifically, while even with one review, a correct acceptance threshold can induce the author to perfectly self-select and only submit papers of positive quality, this threshold may be very high. As a result, it may take many rounds of resubmission to accept the desirable papers. More reviews per round can be interpreted as a more accurate signal, allowing earlier acceptance for desirable papers. We empirically study this tradeoff and find the optimal m.



(a) Binary.



(b) Categorical.

Fig. 5. The review burden vs. review quality under different numbers of solicited reviews in (a) the $(\alpha = 1, \beta, m, V = 5, \eta = 0.7)$-*binary model* and (b) the $(\lambda_R, m, V = 5, \eta = 0.7)$ − $ICLR^{2020,4}$ *noiseless-author model.*

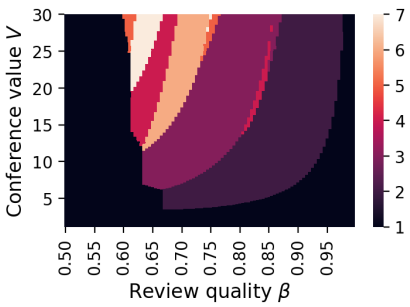
In Fig. 5, we show examples of how m and the review quality affect the review burden. We do not trace out all the points of the QB-tradeoff. Instead, we set the de facto threshold as permissibly as possible while maximizing conference quality. Note that in Fig. 5, the optimal number of reviews per round is either 1 or 2.

Next, in Fig. 6, we investigate how the optimal number of solicited reviews and the review burden depend on the combination of the conference's value V and the review quality, in both the binary model and the categorical models. We observe that, in general, a relatively small number of solicited reviews ($m \leq 3$) is optimal in most cases, especially when the conference is not much more valuable than the sure bet.

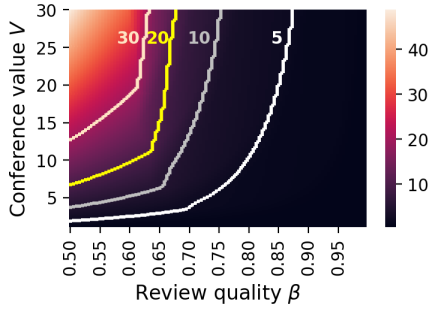
Inspecting the results for the binary model in Fig. 6(a), we observe that a single review is optimal both when the review quality is low (β is close to 0.5) and almost perfect (β is close to 1). The intuition is that when the review quality is poor, slightly more reviews do not help with distinguishing the papers significantly better; on the other hand, when the reviews are almost perfect, the conference does not need more reviews to distinguish the papers. Additional reviews are most beneficial in the intermediate range of β.

B.5 The Tradeoff between Review Quality and Quantity

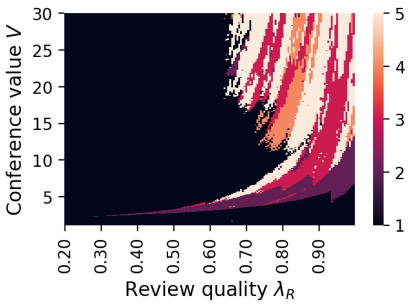
We next consider how a change in the quality of reviews affects the review burden. More specifically, We investigate how many fewer reviews are sufficient to retain the same conference quality when each review has slightly higher quality. Naturally, producing or procuring higher-quality reviews



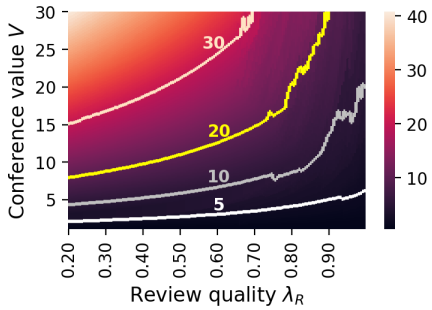
(a) Optimal m, binary.



(b) Review burden, binary.



(c) Optimal m, categorical.



(d) Review burden, categorical.

Fig. 6. The m that leads to the minimum review burden while guaranteeing the maximum conference quality in (a) the $(\alpha = 1, \beta, m, V, \eta = 0.7)$-*binary model* and (c) the $(\lambda_R, m, V, \eta = 0.7)$-*ICLR*^{2020,4} *noiseless-author model*. (b) and (d) show the minimum review burden that can be achieved while guaranteeing the maximum conference quality in the same two models, respectively.

takes extra effort; our results can be interpreted as guidance on whether the extra effort is worth the improvement within the context.

Revisiting Fig. 5, our results show that in general, improving the review quality lowers the review burden. This is true both if we fix m or if we optimize over the best m (which traces out the lower envelope of the plots). For the more controlled binary model (in which review noise can be unambiguously quantified), monotonicity indeed holds throughout. In contrast, in the more complicated categorical model, the review burden R is not always monotone decreasing in the review quality λ_R. This may be in part due to the definition of noisy reviews in the categorical model: the specific interplay between the uniformly random reviews we consider as noise and the learned distribution may lead to such unexpected behavior.

We next investigate the dependency more quantitatively in the binary model in Fig. 5(a) for a fixed value of m. For example, observe that for $m = 3$, the review burden for $\beta = 0.7$ is $R = 3.4$ while for $\beta = 0.6$, it is just shy of 6.7, almost twice as high. Thus, the review burden can be nearly halved by obtaining slightly more accurate reviews. Note that with $m = 3$, the smallest possible review burden is 1.5. This is because when the reviews are perfect, the bad half of the papers will not be submitted, while the good half will be accepted after one round (entailing 3 reviews).

We observe that generally, the effect of improving review quality on the review burden is more significant when the review quality is low. To see this, we look at the lower envelope of Fig. 5(a)

when m is optimized. Here we see that the impact on review burden for improving β is larger when $\beta < 0.75$ than when $\beta > 0.75$. However, in the categorical case in Fig. 5(a), the behavior is different. Apart from $m = 1$, the tradeoff is nearly linear. When looking at the lower envelope, there is a sharp decline in the review burden around $\lambda_R = 0.6$, which could again be caused by noise model idiosyncrasies.

We also consider the joint dependency on the conference's value and the review quality. Fig. 6(a) shows the review burden for different conference values V, optimized over m and the policy threshold. We see that the review burden increases with V; the increase is steeper when the reviews have low quality.

C AUTHORS WITH NOISY SIGNALS: ABM EXPERIMENTS

So far, we have considered the impact of parameters and policies on the conference's tradeoffs in a simple, clean, and analytically tractable model (in Section 4) and in two discrete models, one learned from ICLR data (in Appendix B); there, no clean closed form was available, but analytical results still lent themselves to easy plots. The primary reason for why these models were tractable was that authors had perfect signals about their own papers' qualities, and thus did not update their beliefs in response to reviews.

When authors receive noisy signals themselves, they *will* update their beliefs about their papers' quality based on the reviews they receive. For example, an author who initially received a signal indicating that her paper was of high quality may revise this estimate downward after receiving several negative reviews. As a result, authors may not make the same decision in each iteration; while it may initially be utility-maximizing to submit the paper, after several negative reviews, the author may instead submit to the sure bet.

The required belief updating rules for aggregating the information from multiple conditionally independent signals is unfortunately analytically quite intractable. However, understanding the behavior of more realistic authors (with imperfect information) is also an important robustness check on our results.

In this section, we focus on authors with noisy signals. Due to the aforementioned analytical intractability, we use agent-based simulations (agent-based modeling (ABM)) to evaluate the impact of conference policies on outcomes.

C.1 Experimental Setup

In our ABM experiments, we simulate the submission-review process for n papers. As a brief recap of our model from Section 2, we conceptualize the submission-review process in three phases: submission–reviewing–decision.

In the first phase, each author updates her belief about her paper based on her private signal that is generated based on α, and the reviews from the previous rounds (if any). Given the belief, each author reasons about the expected utility of submitting to the conference or to the "sure bet" option. We consider two strategies for the authors' decision making. The *optimal strategy* assumes that authors are best-responding to the conference's policy. That is, given that the game lasts for T rounds[16], the author uses backward induction with dynamic programming to optimize the decision. Specifically, she first reasons about the expected utility in round T given all possible histories of reviews; she then similarly updates the reviews-utility mapping in rounds $T-1$, $T-2$ and so on. Eventually, she can infer the optimal action in the current round.

We also consider a simpler *myopic strategy*. Here, the author reasons about the expected utility of submitting to the conference by assuming that after a rejection, she will submit to the sure bet

[16]Due to running time concerns, we set $T = 10$ in our experiments.

option. That is, a myopic author in round t does not foresee the future after round $t+1$. Empirically, this is nearly the same as the optimal strategy. The myopic strategy asks: is it better to submit one more round before giving up or to give up now? However, it can be the case that submitting two more rounds before giving up is better than giving up now which in turn is better than giving up after one round of submission. However, such cases are quite rare when the same acceptance policy threshold is used in each round.

In the second phase, the conference obtains m reviews for each submission. These reviews are drawn i.i.d. from the review signal distribution $\boldsymbol{\beta}$, conditioned on the ground truth quality of the submission.

In the third phase, for each submission, the conference makes a decision of acceptance or rejection based on a threshold policy with threshold τ. Given the m i.i.d. reviews sampled in the second phase, the conference can infer the conditional expected quality of the submission and accept or reject the paper based on the threshold policy (described in Appendix B.2).

The three phases are repeated for T rounds. After T rounds, all papers that have not been accepted (by either the top conference or the sure bet) are submitted to and accepted by the sure bet.

Choosing model parameters. We consider the same two types of model as in Appendix B: a simple binary quality model (which allows us to characterize model properties concisely with few parameters), and the more fine-grained categorical models learned from ICLR data.

For the binary model, we assume that the authors use the optimal strategy. When studying the categorical model, we assume that the authors apply the myopic strategy; this is done due to running-time concerns, because finding the optimal strategy would involve search over too many possible options.

For both models, we fix $m = 3$, $V = 3$ and $\eta = 0.7$. There are $n = 10000$ submissions, and we consider $T = 10$ rounds. We note that in most cases, $T = 10$ is enough for the conference to accept all submitted papers.

C.2 QB-Tradeoff and Acceptance Rate

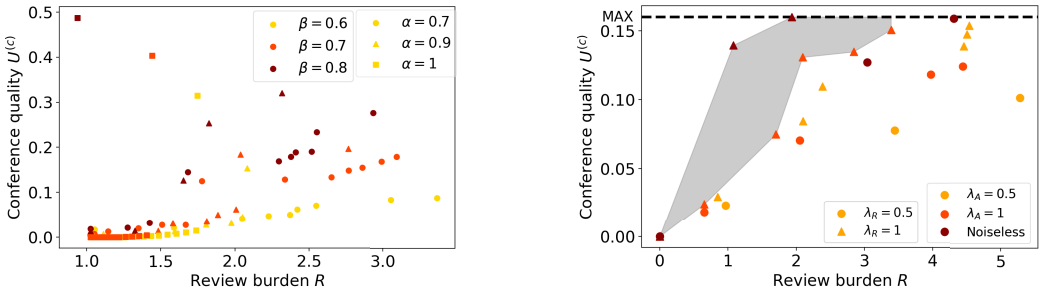
In Fig. 7, we show the QB-tradeoff and its dependence on the signal qualities of the authors and reviewers. In Fig. 7(b), the parameters λ_A and λ_R with which the author and reviewer, respectively, obtain a meaningful (as opposed to uniformly random) signal captures the signal quality; see the definition in Appendices B.1 and E.

C.2.1 The QB-tradeoff. We first note that the Pareto frontier of the QB-tradeoff consists of a larger number of points when authors are noisy. For example, in the binary case, the acceptance threshold can be set so that agents with signals indicating high-quality papers only submit once (or alternatively twice) before giving up, while agents with signals indicating low-quality papers never submit. In the case where agents submit only once before giving up, the review burden will be less, but so will the conference quality, as some high-quality submissions will give up before being accepted. Cases like this can never occur when agents have perfect signals about their papers.

Then, an immediate observation in Fig. 7 is that improving the signal quality for either the author or the reviewers is beneficial to the quality-review tradeoff.

In Fig. 7(b), we hypothesize that the Pareto frontier of the current review system is located in the grey area between the brown triangles and the red triangles. The former represents the case of noiseless authors, and the latter represents the case of authors whose signals are as noisy as the reviewers'. In both cases, the reviewers' signals are learned from the ICLR 2020 dateset. Furthermore, the current review system can achieve at least 94% of the maximum conference quality with at most 180% of the review burden achieved if the authors were noiseless. This implies that 1)

high quality can still be achieved at the cost of a higher review burden; 2) improving the review quality or the authors' signals can substantially mitigate the review burden.



(a) Binary.

(b) Categorical.

Fig. 7. Pareto optimal points of the quality-review tradeoff under different review qualities and author signal qualities, resulting from different acceptance thresholds in (a) the $(\alpha, \beta, m = 3, V = 5, \eta = 0.7)$-*binary model* and (b) the $(\lambda_A, \lambda_R, m = 3, V = 5, \eta = 0.7)$-*ICLR*^{2020,4} *model*. In (a), authors are best-responding; and in (b), authors use the myopic strategy if noisy and are best-responding if noiseless. The grey area is defined by the Pareto frontier of the red triangles and the brown triangles; we hypothesize that this is the Pareto frontier of the real review system.

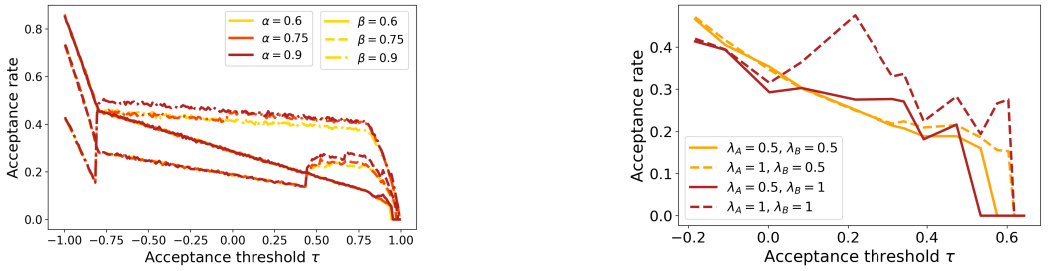
C.2.2 The Acceptance Rate. As in Fig. 4(b), the acceptance rate as a function of the acceptance threshold — shown in Fig. 8 — roughly follows a sawtooth shape: as the threshold increases, the rate decreases, except at discrete points. At these points, the lowest remaining quality level of papers is not submitted any more, leading to a sudden jump in acceptance rate. We can also observe that this effect is more pronounced when authors have better signals about their papers' qualities.

We further look at the effects of the author/review signals on the acceptance rate. On one hand, Fig. 8 shows that better author self-selection quality leads to a higher acceptance rate. The effect of better self-selection emerges after at least the bottom quality tier of papers stops submitting. On the other hand, the effect of the review quality can be discussed in three cases. First, when the acceptance threshold is low, e.g., lower than 0.05 in Fig. 8(b), higher review quality decreases the acceptance rate, since more bad papers are rejected. Second, in the middle case ($0.05 < \tau < 0.45$), higher review quality generally increases the acceptance rate since it helps the authors' self-selection. Finally, when the threshold is very high, e.g., higher than 0.45, higher review quality may decrease the acceptance rate if the authors' signals are rather noisy. (The red solid line is below the yellow solid line.) This is because when authors are noisy, even authors with good signals become less confident about their papers' prospects when the acceptance threshold is strict. Thus, compared with the case when the review quality is low, those authors may want to quit earlier when the acceptance threshold increases.

C.3 Review Quality vs. Quantity

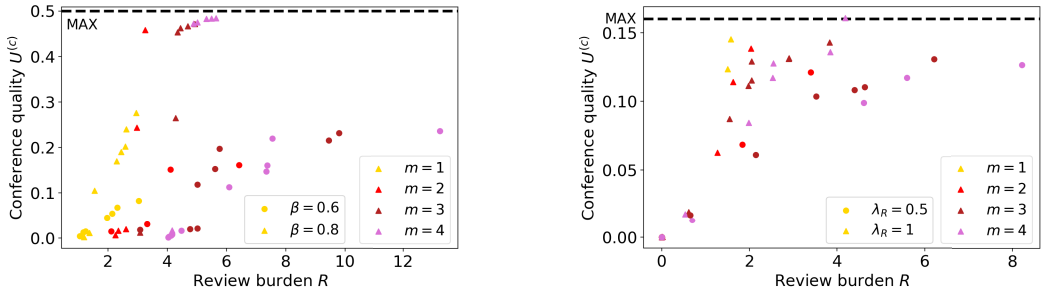
We next investigate the tradeoff between the number and quality of reviews, in the presence of noisy authors.

Fig. 9 is akin to Fig. 7; instead of varying the author's noise, we varied the number m of reviews. In general, a larger number of solicited reviews m leads to a heavier review burden, but can achieve a slightly higher maximum conference quality. The intuition is that more reviews per round helps the authors update their belief about their papers faster. Thus, better self-selection can be induced



(a) Binary. (b) Categorical.

Fig. 8. Acceptance rate as a function of the acceptance threshold in (a) the $(\alpha, \beta, m = 3, V = 5, \eta = 0.7)$-*binary model* and (b) the $(\lambda_A, \lambda_R, m = 3, V = 5, \eta = 0.7)$ − *ICLR*^{2020,4} *model*. In (a), authors are best-responding; and in (b), authors using the myopic strategy.

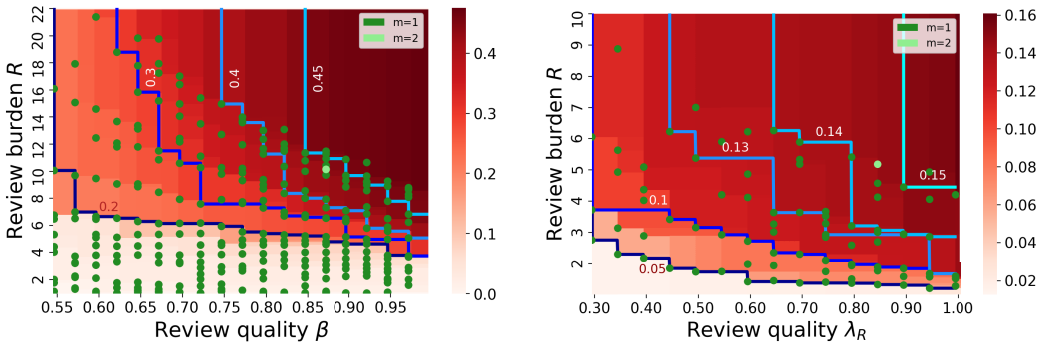


(a) Binary. (b) Categorical.

Fig. 9. Pareto optimal points of the quality-review tradeoff under different review qualities and m, resulting from different acceptance thresholds in (a) the $(\alpha = 0.8, \beta, m, V = 5, \eta = 0.7)$-*binary model* and (b) the $(\lambda_A = 1, \lambda_R, m, V = 5, \eta = 0.7)$ − *ICLR*^{2020,4} *model*. In (a), authors are best-responding; and in (b), authors are using the myopic strategy (Appendix C.1) if noisy and are best-responding if noiseless.

in fewer rounds which results in fewer mistakes made by the conference. Again, in line with Fig. 5, our results show that the Pareto frontiers when the number of solicited reviews is larger than 3 are largely dominated by the frontiers when $m \leq 3$. However, we observe that the disadvantage of using $m = 1$ is enlarged when authors are noisy: authors cannot learn the true quality of their papers fast enough and because we run the resubmission process for $T = 10$ rounds (due to computation issues), this results in significantly lower maximal conference quality. Even worse, in Fig. 9(b), the conference cannot obtain positive conference quality with $m = 1$ when reviews are rather noisy. (There is only one yellow dot at (0,0).)

Fig. 10 provides a different visualization of the interaction of the conference quality and review burden, which now also incorporates the review quality. For each review quality, we compute the maximum conference quality attainable without going over some review burden. We see that higher review quality yields more favorable QB-tradeoffs. We also note that nearly all Pareto frontier points correspond to $m = 1$ reviews, although a few correspond to $m = 2$ reviews. This is in line with Fig. 6(a), where the optimal number of solicited reviews is $m = 1$ when $V = 5$.



(a) Binary.

(b) Categorical.

Fig. 10. The maximum conference utility that can be achieved with review quality β (or λ_R) and review burden R when authors are noisy in (a) the $(\alpha = 0.8, \beta, m, V = 5, \eta = 0.7)$-*binary model* and (b) the $(\lambda_A = 1, \lambda_R, m, V = 5, \eta = 0.7)$ − *ICLR*^{2020,4} *model*. For each β and λ_R, we search over all numbers of solicited reviews $m \in \{1, \ldots, 5\}$, and for each m, we search over acceptance thresholds τ from $[\min(Q), \max(Q)]$ with step size 0.01. Then, the Pareto optimal points are plotted as green dots, and the heat map is drawn to represent the maximum conference quality that can be achieved with the particular review quality and review burden. (The maximum conference quality that can possibly be reached is 0.5 in (a) and 0.161 in (b).)

D INSTITUTIONAL MEMORY

Most conferences' acceptance policies are memoryless, in the sense that resubmissions are treated the same as new submissions. However, in part to deal with the large number of papers that are repeatedly resubmitted, several conferences have experimented with models that have "institutional memory." We consider the following types of policies which are not memoryless and contain some institutional memory.

Time Limited, Fixed Threshold: The simplest way to incorporate memory into the submission process is to limit the number of times the same paper can be submitted. We call such a policy a *T-round fixed-threshold policy*.

Time Limited, Variable Threshold: A generalization of T-round fixed threshold policies is to allow different acceptance thresholds for different rounds. This allows a conference to set higher/lower standards for resubmissions. However, we require the conference to solicit the same number of reviews for each round.[17] Formally, a *round-dependent threshold policy* is defined by a threshold vector $\boldsymbol{\tau} = \left[\tau^{(1)}, \tau^{(2)}, \ldots, \tau^{(T)}\right]$; in round $t \leq T$, a paper with reviews $\boldsymbol{s}$ is accepted if and only if its expected quality conditioned on $\boldsymbol{s}$ is at least $\tau^{(t)}$.

Review Following: Under a *T-round review-following threshold memory policy*, not only does the conference track the number of resubmissions, but also considers all past reviews as equal (additional) reviews of resubmissions; that is, reviews are treated identically regardless

[17] We do so for two reasons. First, this reduces the policy space — this is significant in terms of computation when optimizing over policies with memory. Second, it excludes highly unrealistic policies with very specific dependency on model parameters. For example, when authors are noiseless, applying a policy with memory can achieve maximum conference quality with minimum review burden: the conference rejects all submissions $T - 1$ times without review. In round T, one review is solicited, and the paper is accepted if and only if the expected quality conditioned on the review is larger than τ; finally, in round $T + 1$, the submission is accepted without review. A careful choice of T and τ, taking advantage of authors' patience (or lack thereof) and knowledge of their own paper's quality, ensures that no negative-utility papers are submitted, yet all positive-utility papers are submitted and eventually accepted.

of which round they were provided in.[18] Again, we have the conference obtain the same number of reviews in each round of resubmissions, i.e., $m_t = m$ for all t. The conference commits to a number T of rounds and rejects any paper that has been submitted T times. The conference also commits to a sequence of thresholds $\boldsymbol{\tau} = \left(\tau^{(1)}, \tau^{(2)}, \ldots, \tau^{(T)}\right)$, such that in round t, a paper with historical reviews $\boldsymbol{c}_t$ is accepted if and only if its expected quality conditioned on $\boldsymbol{c}_t$ is at least $\tau^{(t)}$.

All three policy types are time limited, in that the number of times any particular paper can be submitted is capped. This is similar to certain National Science Foundation programs (e.g., CAREER), where the number of times a proposal (or sometimes author) can submit is limited. Recall that in **??**, we discussed the time-limited, fixed-threshold policy and showed that the authors' best response remains the same as in the memoryless case. In the following section, we further analytically investigate this review policy with a focus on how it affects the QB-tradeoff.

In addition, we consider two policies that allow different thresholds in different rounds in two different ways: in time-limited, variable-threshold policies, the decision only depends on the most recent reviews, but the threshold may change depending on how many times the paper has already been submitted; in review following policies, the acceptance policy is applied to all prior reviews. This models the increasingly common policy of requiring resubmissions to be accompanied by previous reviews (e.g., at IJCAI).

These two generalizations are more complicated to analyze, and we use ABMs to investigate their QB-tradeoff.

Perhaps subtly, the review-following policy does not fully subsume the other two policies, because past reviews cannot be treated differently from new reviews. For example, the review-following policy cannot simulate a time-limited fixed threshold policy in which every round, a paper obtains two reviews and is accepted iff both reviews are positive. The reason is that the review-following policy cannot distinguish the case of having one positive review in each round (which should lead to rejection) from the case of having zero positive reviews in the first round and two positive reviews in the second round (which should lead to acceptance).

D.1 Time-limited Policy With Fixed Threshold

We start with this simple policy where the conference merely restricts the number of times a paper can be submitted (to some value T). We will show that under such a policy, noiseless authors will respond exactly as if the conference allowed unlimited resubmissions.

Proposition D.1. *Consider a conference which allows a paper to be submitted at most T times, and for each of these submissions independently decides whether to accept the paper, according to the same monotone policy ϕ. Then, the author's best response is to submit the paper (in each round) if and only if $P_{acc}(\phi, Q) \geq \rho$, and to submit to the sure bet option otherwise.*

Proof. The author will submit a paper with quality $Q = q$ iff her expected utility is at least 1. We compute the expected utility for an author who submits the paper exactly T times, akin to the proof of Proposition 3.1:

[18] As such, they do not serve the purpose of verifying whether authors addressed concerns about previous versions of their paper.

$$
\begin{aligned}
U^{(a)}(\phi, q) &= V \cdot \sum_{t=1}^{T} \left(P_{\text{acc}}(\phi, q) \cdot (\eta \cdot (1 - P_{\text{acc}}(\phi, q)))^{t-1}\right) + (\eta \cdot (1 - P_{\text{acc}}(\phi, q)))^{T} \\
&= V \cdot \frac{P_{\text{acc}}(\phi, q) \cdot \left(1 - (\eta \cdot (1 - P_{\text{acc}}(\phi, q)))^{T}\right)}{1 - \eta \cdot (1 - P_{\text{acc}}(\phi, q))} + (\eta \cdot (1 - P_{\text{acc}}(\phi, q)))^{T} \\
&= \left(\frac{V \cdot P_{\text{acc}}(\phi, q)}{1 - \eta \cdot (1 - P_{\text{acc}}(\phi, q))} - 1\right) \cdot \left(1 - (\eta \cdot (1 - P_{\text{acc}}(\phi, q)))^{T}\right) + 1.
\end{aligned}
$$

To determine when this utility is at least 1 (the utility the author would obtain from choosing the sure bet option right away), we need to determine when the product of the first two terms is non-negative. The second factor is always positive, and the first is non-negative if and only if $P_{\text{acc}}(\phi, q) \geq \rho$. This completes the proof. □

Notice that while the author's expected *utility* depends on T, her best response does not. Hence, as with an unlimited number of submissions, an author will either submit forever or not even submit once, and the author's threshold for doing so is the same as with an infinite number of resubmissions.

We next investigate how the restricted number of resubmissions, T, affects the QB-tradeoff. Proposition D.1 allows us to characterize the expected value of the conference quality and the review burden. For a given acceptance threshold τ, let $\mathcal{S}_\tau \subseteq \mathcal{Q}$ be the set of paper qualities which an author will submit at this threshold. When $\mathcal{Q}$ is discrete[19], the expected conference quality and review burden are as follows:

$$
\begin{aligned}
U^{(c)}(\tau) &= \sum_{q \in \mathcal{S}_\tau} p(q) \cdot q \cdot \left(P_{\text{acc}}(\tau, q) + (1 - P_{\text{acc}}(\tau, q))P_{\text{acc}}(\tau, q) + \cdots + (1 - P_{\text{acc}}(\tau, q))^{T-1} P_{\text{acc}}(\tau, q)\right) \\
&= \sum_{q \in \mathcal{S}_\tau} p(q) \cdot q \cdot \left(1 - (1 - P_{\text{acc}}(\tau, q))^{T}\right) \\
R(\tau) &= m \cdot \sum_{q \in \mathcal{S}_\tau} p(q) \cdot \left(1 + (1 - P_{\text{acc}}(\tau, q)) + \cdots + (1 - P_{\text{acc}}(\tau, q))^{T-1}\right) \\
&= m \cdot \sum_{q \in \mathcal{S}_\tau} p(q) \cdot \frac{1}{P_{\text{acc}}(\tau, q)} \cdot \left(1 - (1 - P_{\text{acc}}(\tau, q))^{T}\right).
\end{aligned}
$$

Fig. 11 shows the QB-tradeoff for the time-limited fixed-threshold policies in the continuous model. Not surprisingly, the conference can reduce the review burden at the expense of quality by lowering T. Doing so reduces the maximum conference quality that can be reached by the review policy but can also reduce the review burden if the desired conference quality is reachable.

Fig. 12 shows results analogous to Fig. 11, but in the categorical model, and further varying the numbers of solicited reviews and the acceptance threshold (to accept different categories of paper qualities). The same pattern can also be observed here: by comparing different colors of dots while fixing an m (fixing a line), we observe that the conference can reduce the review burden at the expense of quality by lowering T.

As a second observation, when T is small, a large number of reviews per submission contributes more to the conference quality. This can be seen by fixing a color and a shape of marker and looking at the dots on different types of the lines: solid lines (one review) result in the lowest conference

[19]For continuous qualities, the corresponding quantities are obtained by replacing sums by integrals and probabilities by densities.

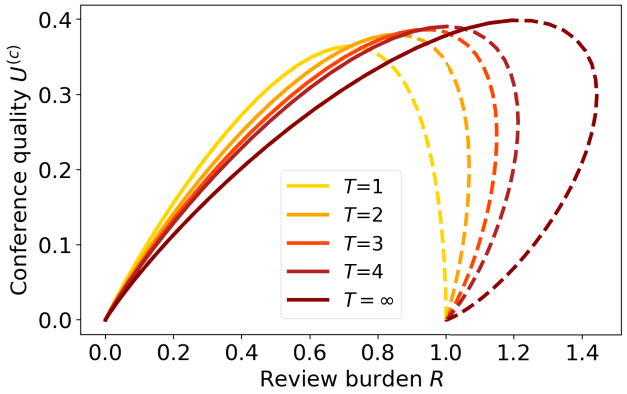


Fig. 11. The Pareto optimal curves of the QB-tradeoff under T-round fixed threshold policies in the $(\sigma = 0.5, \mu_{\boldsymbol{p}} = 0, \sigma_{\boldsymbol{p}} = 1, m = 1, V = 3, \eta = .7)$-*Double Gaussian noiseless-author model.* Pareto dominated points are shown as dashed lines, while undominated points are shown as solid lines.

quality when T is small. The intuition is that when m is small, the conference has to apply a very strict acceptance threshold to discourage low-quality papers from being submitted. As a result, such a policy will need a large number of rounds to accept the good papers. We conclude that if the conference severely limits the number of times a paper can be submitted, soliciting more reviews per paper will contribute more to a higher conference quality at the expense of additional reviews.

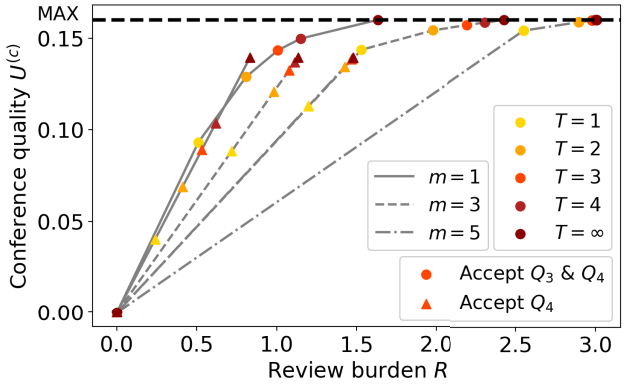


Fig. 12. The Pareto optimal points of the quality-burden tradeoff under T-round fixed threshold policies in the $(\lambda_R = 1, m, V = 5, \eta = 0.7)$-$ICLR^{2020,4}$ *noiseless-author model.* The colors of the markers distinguish T, and the shapes of the markers denote the accepted papers' qualities with $Q_1 < Q_2 < 0 < Q_3 < Q_4$. Furthermore, we vary the number m of reviews per paper, shown by the line type.

D.2 More Fine-Grained Memory, and Noisy Authors

Next, we investigate the extent to which more fine-grained memory — different acceptance thresholds in different rounds and reuse of past reviews of a paper — may further improve the quality-review tradeoff for the conference. We do so with noisy authors, and therefore — as before — use ABM experiments instead of a theoretical analysis.

D.2.1 Experimental Setup. Our approach involves searching over policies with T submissions per paper. Unfortunately, the number of such policies grows exponentially in T. We therefore restrict our experiments to $T = 5$, and set $m = 3$ for only the binary model. Computing the optimal strategy for categorical or continuous models requires solving a dynamic program with a much larger state space; the resulting computational requirements prevent us from including such experiments.

We generate candidate policies by sampling thresholds. For T-round fixed-threshold policies, we select 40 candidate thresholds τ from the set $\{-1, -0.95, \ldots, 0.95\}$; in each run, one of these thresholds is used for all rounds. For the other two types of policies, to reduce the number of samples, we only sample the threshold $\tau^{(t)}$ for the first three rounds while fixing $\tau^{(t)}$ for $t = 4, 5$. We sample 40^3 candidate threshold vectors $\boldsymbol{\tau}$, as follows: for each $t \in \{1, 2, 3\}$, we select 40 thresholds $\tau^{(t)}$ from $\{-1, -0.95, \ldots, 0.95\}$ which gives us 40^3 vectors of length 3. For each $t \in \{4, 5\}$, we fix $\tau^{(t)}$ such that the paper is accepted if and only if the newly sampled review in round t is 1. (Note that the review signal in the binary model is either -1 or 1.)

D.2.2 Results. Figure 13 shows the Pareto optimal points of the QB-tradeoff for each of the three review policies. We summarize the main takeaways as follows:

- Compared with the time-limited fixed-threshold policy, round-dependent threshold policies (the variable-threshold policy and the review-following policy) have slightly higher maximum conference quality. Furthermore, in general, for each of the Pareto optimal points of the former, there exists a Pareto optimal point of the latter which dominates it. Our results imply that having historical reviews following the submissions can indeed help reduce the review burden and thus improve the QB-tradeoff, though it's effect is rather marginal especially in terms of improving the conference's quality.
- We further observe that for both the variable-threshold policy and the review-following policy, with round-dependent thresholds, the Pareto optimal thresholds that lead to large conference qualities tend to have the following pattern: review papers strictly in the first two rounds and more leniently after that.

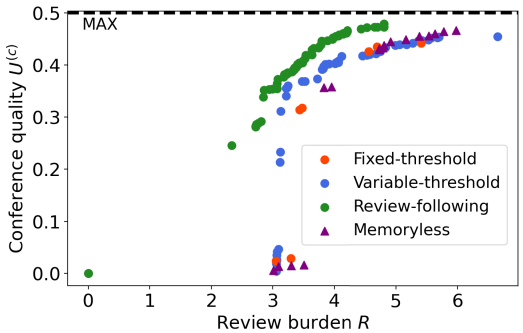


Fig. 13. The Pareto optimal points of the quality-burden tradeoff for three types of policy when authors are noisy in the $(\alpha = 0.75, \beta = 0.75, m = 3, V = 5, \eta = 0.5, T = 5)$-*Binary model. The Pareto optimal points of the memoryless policy are also shown for comparison; the memoryless policy is approximated by setting a large* $T = 50$.

E LEARNING PARAMETERS OF THE CATEGORICAL MODEL FROM DATA

We set the parameters of our model based on the OpenReview datasets of submissions and reviews for ICLR 2020 [1] and ICLR 2021 [2]. The datasets contain about 1500 and 2500 submissions,

respectively; typically, each submission is reviewed thrice, with scores in $\Sigma = \{0, 1, \ldots, 9\}$. We apply the same learning algorithm to each of the two datasets separately, yielding two plausible parameter settings for evaluation.

The Number of Paper Quality Scores and Signals. While the set $\Sigma = \{0, 1, \ldots, 9\}$ of available review scores is known, the number (or set) of different paper qualities is not. Thus, our goal is to simultaneously learn the number of paper qualities, the paper quality distribution, and the distribution of reviewer signals conditioned on the paper's quality.

To do so, we exhaustively try all numbers L of paper quality scores in $\{2, \ldots, 10\}$; for each, we apply a variant of the EM algorithm described below. Once the EM algorithm has converged, we evaluate the likelihood of the learned model for the held-out test data, and retain the model(s) with the highest likelihood scores.

Given a choice of L, to learn the paper quality distribution $\boldsymbol{p}$ and conditional review distribution $\boldsymbol{\beta}$, we apply the EM algorithm with regularization and cross-validation to avoid overfitting [9].

Specifically, by cross-validation, we first randomly divide the dataset into five subsets with approximately equal size. We choose one of the subsets as the test set while the remaining 80% of data form the training set. This step is repeated for five times: a different one of the five subsets is used as the test set. Given the training and test dataset, for each $L \in \{2, 3, \ldots, 10\}$, we run the EM algorithm for 100 iterations on the training set to estimate the quality of each paper and the confusion matrix for reviewers (i.e., the matrix of review score probabilities conditional on ground truth quality); the EM algorithm alternates between updating the quality distribution with fixed confusion matrix, and updating the confusion matrix with fixed quality distribution. We apply regularization every M steps: given an estimated confusion matrix $\boldsymbol{\beta}^{(k)}$ after the k-th iteration, we perturb $\boldsymbol{\beta}^{(k)}$ with a small amount of noise so that each row i becomes the convex combination $0.99 \cdot \boldsymbol{\beta}_i^{(k)} + 0.01 \cdot \frac{1}{|\Sigma|}\mathbf{1}$; here, $\frac{1}{|\Sigma|}\mathbf{1} \in \mathbb{R}^{|\Sigma|}$ is the uniform distribution on signals. After the EM algorithms converges, we evaluate the likelihood of the test data given the trained model, i.e. $\boldsymbol{p}$ and $\boldsymbol{\beta}$.

For each value of L, we judge the learned model based on the likelihood averaged over the five times cross-validation. The paper quality space size L that has the maximum averaged likelihood is selected. Finally, we output the paper quality distribution $\boldsymbol{p}$ as well as the confusion matrix $\boldsymbol{\beta}$ as the average of the learned parameters for each of the five runs with different test sets. We find that $L \in \{4, 5\}$ tends to fit the data well in both of the datasets. For robustness, we conduct our experiments for both datasets for $L = 4$ and 5. The resulting parameters are shown in Table 1. For experiments in which the authors also receive only noisy signals (instead of the ground truth quality), we set the confusion matrix for the authors $\boldsymbol{\alpha}$ to be the same as the one for reviewers, $\boldsymbol{\beta}$; this is because unfortunately, very little data are available that show how authors evaluate their own papers.

In some of our experiments, we want to explicitly evaluate the impact of increasing the noise in reviews. To do so, we consider reviewer signal matrices which are a convex combination of the learned signal matrix $\boldsymbol{\beta}$ with the uniform signal distribution $\frac{1}{|\Sigma|}\mathbf{1}$; this corresponds to a reviewer who assigns a uniformly random score with probability $1 - \lambda_R$. The weight $\lambda_R \in [0, 1]$ placed on the learned distribution $\boldsymbol{\beta}$ then captures the quality of the signal. Similarly, λ_A controls the weight of the confusion matrix of the authors' signal.

The Value of Paper Quality. First, for each paper i in the dataset, we take the average over the review scores, denoted as $\bar{s}_i^{(r)}$. Then, we set the value of the quality of paper i as $\psi(\bar{s}_i^{(r)})$, where $\psi : [0, 9] \to \mathbb{R}$ is an increasing function that maps from the average score to the quality of

a paper. In our experiments, we set ψ as a reversed (and shifted and scaled) sigmoid function, i.e. $\psi(x) = \log \frac{x+0.01}{9.01-x}$ for $x \in [0, 9]$.

Next, the task is to infer the a label for each paper i that is a probability distribution of a paper i' quality given $\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{\beta}}$ and $s_j^{(r)}$, i.e. $l_{i,k} = \Pr[Q = q_k | s_i^{(r)}, \boldsymbol{p}, \boldsymbol{\beta}]$ for $1 \le k \le L$. This probability can be computed based on Bayes' rule.

Finally, we can set the value of paper quality as a weighted average

$$q_k = \frac{\sum_i^n l_{i,k} \cdot \psi\left(\bar{s}_i^{(r)}\right)}{\sum_i^n l_{i,k}}. \tag{4}$$

E.1 The Learned Parameters

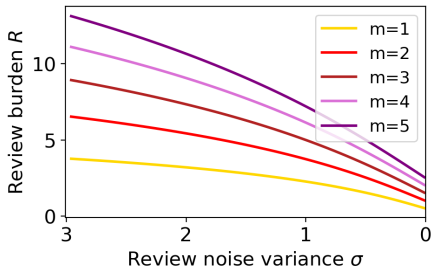
Table 1 summarizes our parameters for two datasets. For each dataset, we infer parameters for two sizes $L = 4, 5$ of the set of paper qualities.

	$\boldsymbol{p}$	$\boldsymbol{\beta}$	Q
ICLR 2020, $L = 4$	$\begin{bmatrix} 0.0772 \\ 0.3987 \\ 0.2648 \\ 0.2593 \end{bmatrix}$	$\begin{bmatrix} 0.0400 & 0.1706 & 0.4200 & 0.2729 & 0.0685 & 0.0028 & 0.0247 & 0.0002 & 0.0002 & 0.0001 \\ 0.0004 & 0.0123 & 0.1195 & 0.3816 & 0.2948 & 0.1288 & 0.0450 & 0.0119 & 0.0056 & 0.0001 \\ 0.0001 & 0.0043 & 0.0226 & 0.0959 & 0.3235 & 0.3626 & 0.1648 & 0.0188 & 0.0073 & 0.0001 \\ 0.0001 & 0.0016 & 0.0090 & 0.0302 & 0.0534 & 0.2922 & 0.4285 & 0.1427 & 0.0375 & 0.0048 \end{bmatrix}$	$\begin{bmatrix} -1.1145 \\ -0.4079 \\ 0.0544 \\ 0.5606 \end{bmatrix}$
ICLR 2020, $L = 5$	$\begin{bmatrix} 0.0923 \\ 0.282 \\ 0.25 \\ 0.206 \\ 0.1696 \end{bmatrix}$	$\begin{bmatrix} 0.0340 & 0.1538 & 0.3981 & 0.2865 & 0.0928 & 0.0085 & 0.0257 & 0.0002 & 0.0003 & 0.0001 \\ 0.0003 & 0.0077 & 0.1214 & 0.4345 & 0.2495 & 0.1266 & 0.0415 & 0.0153 & 0.0031 & 0.0001 \\ 0.0001 & 0.0079 & 0.0523 & 0.1603 & 0.4388 & 0.2270 & 0.0918 & 0.0184 & 0.0033 & 0.0001 \\ 0.0001 & 0.0031 & 0.0126 & 0.0668 & 0.1180 & 0.4503 & 0.3010 & 0.0444 & 0.0034 & 0.0003 \\ 0.0001 & 0.0014 & 0.0082 & 0.0206 & 0.0604 & 0.2216 & 0.4446 & 0.1823 & 0.0535 & 0.0073 \end{bmatrix}$	$\begin{bmatrix} -1.0552 \\ -0.4337 \\ -0.1378 \\ 0.2747 \\ 0.6454 \end{bmatrix}$
ICLR 2021, $L = 4$	$\begin{bmatrix} 0.0267 \\ 0.287 \\ 0.6152 \\ 0.0711 \end{bmatrix}$	$\begin{bmatrix} 0.0228 & 0.2288 & 0.4211 & 0.2415 & 0.0639 & 0.0201 & 0.0003 & 0.0010 & 0.0002 & 0.0001 \\ 0.0007 & 0.0241 & 0.1568 & 0.3899 & 0.2712 & 0.1134 & 0.0316 & 0.0056 & 0.0063 & 0.0003 \\ 0.0008 & 0.0062 & 0.0583 & 0.1805 & 0.2783 & 0.2617 & 0.1619 & 0.0450 & 0.0062 & 0.0011 \\ 0.0021 & 0.0031 & 0.0070 & 0.0544 & 0.1173 & 0.2435 & 0.3725 & 0.1256 & 0.0701 & 0.0043 \end{bmatrix}$	$\begin{bmatrix} -1.1544 \\ -0.4886 \\ -0.0331 \\ 0.4927 \end{bmatrix}$
ICLR 2021, $L = 5$	$\begin{bmatrix} 0.0228 \\ 0.2712 \\ 0.306 \\ 0.3304 \\ 0.0695 \end{bmatrix}$	$\begin{bmatrix} 0.0260 & 0.2320 & 0.4249 & 0.2337 & 0.0630 & 0.0183 & 0.0003 & 0.0014 & 0.0002 & 0.0001 \\ 0.0007 & 0.0267 & 0.1616 & 0.3966 & 0.2672 & 0.1074 & 0.0285 & 0.0050 & 0.0060 & 0.0004 \\ 0.0011 & 0.0049 & 0.0600 & 0.1957 & 0.3141 & 0.2614 & 0.1253 & 0.0301 & 0.0063 & 0.0011 \\ 0.0006 & 0.0068 & 0.0578 & 0.1734 & 0.2435 & 0.2578 & 0.1947 & 0.0512 & 0.0128 & 0.0015 \\ 0.0026 & 0.0037 & 0.0038 & 0.0616 & 0.1390 & 0.2640 & 0.3416 & 0.1075 & 0.0718 & 0.0046 \end{bmatrix}$	$\begin{bmatrix} -1.1795 \\ -0.5077 \\ -0.0937 \\ 0.0153 \\ 0.4641 \end{bmatrix}$

Table 1. Parameters learned from the ICLR datasets. The rows of the confusion matrix are ranked based on the expected scores (from low to high). That is, the kth row of the confusion matrix $\boldsymbol{\beta}$ has lower average score than the $(k+1)$-st row. Given this ranking, we observe that the value of paper qualities learned from our method is monotone increasing in k.

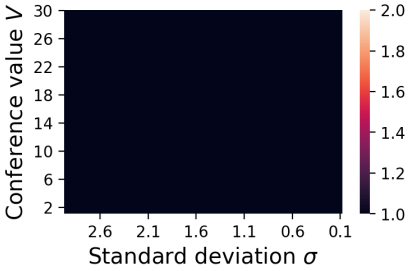
F REVIEW QUALITY AND QUANTITY IN THE CONTINUOUS MODEL

Fig. 14 and Fig. 15 are analogous to Fig. 5 and Fig. 6 for the continuous model, respectively. Again, the results are in line with Appendix B.4 where soliciting one review per paper is optimal for the searched range of parameters. Although this cannot be interpreted as a theorem (one can find extreme cases such that soliciting more than one reviews is optimal), we emphasize that in general, our results in the Gaussian model also suggests a small number of solicited reviews.

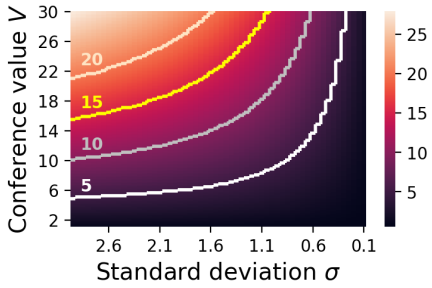


(a) Continuous.

Fig. 14. The review burden vs. review quality under different numbers of solicited reviews in the $(\sigma, \mu_{\boldsymbol{p}} = 0, \sigma_{\boldsymbol{p}} = 1, m, V = 5, \eta = 0.7)$-*Double Gaussian model.*



(a) Optimal m, continuous.



(b) Review burden, continuous.

Fig. 15. (a) The m that leads to the minimum review burden while guaranteeing the maximum conference quality, and (b) the minimum review burden that can be achieved while guaranteeing the maximum conference quality, both are in the $(\sigma, \mu_{\boldsymbol{p}} = 0, \sigma_{\boldsymbol{p}} = 1, m, V, \eta = 0.7)$-*Double Gaussian model.*