
TRUTHFUL AGGREGATION OF LLMs WITH AN APPLICATION TO ONLINE ADVERTISING

Ermis Soumalias
University of Zurich
ETH AI Center
ermis@ifi.uzh.ch

Michael J. Curry
Harvard University
University of Zurich
ETH AI Center
mcurry@seas.harvard.edu

Sven Seuken
University of Zurich
ETH AI Center
seuken@ifi.uzh.ch

ABSTRACT

The next frontier of online advertising is revenue generation from LLM-generated content. We consider a setting where advertisers aim to influence the responses of an LLM to align with their interests, while platforms seek to maximize advertiser value and ensure user satisfaction. The challenge is that advertisers' preferences generally conflict with those of the user, and advertisers may misreport their preferences. To address this, we introduce MOSAIC, an auction mechanism that ensures that truthful reporting is a dominant strategy for advertisers and that aligns the utility of each advertiser with their contribution to social welfare. Importantly, the mechanism operates without LLM fine-tuning or access to model weights and provably converges to the output of the optimally fine-tuned LLM as computational resources increase. Additionally, it can incorporate contextual information about advertisers, which significantly improves social welfare. Through experiments with a publicly available LLM, we show that MOSAIC leads to high advertiser value and platform revenue with low computational overhead. While our motivating application is online advertising, our mechanism can be applied in any setting with monetary transfers, making it a general-purpose solution for truthfully aggregating the preferences of self-interested agents over LLM-generated replies.

1 INTRODUCTION

Large language models (LLMs) are becoming ubiquitous – as coding assistants, as chat interfaces, and even as alternatives to search engines (Bommasani et al., 2022). To ensure their usefulness, it is essential to closely align LLM outputs with user preferences. However, in general, there may be multiple interested parties who disagree on the desired behavior of *the same LLM*. This leads to the question of *how to generate LLM replies given multiple conflicting preferences*.

The motivating application for our work is online advertising, which is the main source of revenue for large tech companies such as Google, Meta, and Twitter. For example, in 2023, Meta's advertising revenue of 132 billion USD was more than 97% of its total revenue (Meta, 2024). Ad auctions are the workhorse mechanism to determine the placement and prices of commercial content (Varian, 2007; Edelman et al., 2007). As platforms are now working on integrating advertisements into LLM-generated content (Criddle, 2024), new auction mechanisms are needed for this application.

We present a new auction mechanism for this problem, which we call *MOSAIC (Mechanism for Optimally Sampling and Aggregating LLM Outputs with Incentive Compatibility)*. The advertisers are paying, not for some particular item or bundle, but rather to influence the output generated by an LLM in a direction closer to their own preferences. Although our mechanism could be of interest whenever one has to aggregate the preferences of multiple self-interested agents over LLM behavior (as long as it is reasonable to charge monetary payments), we see online advertising as the most natural setting of interest. For this reason, we refer to participants in the mechanism as *advertisers*.

1.1 PROBLEM SETTING

Consider a situation where a user queries an LLM with a specific question. We assume the following:

-
- There is a *reference LLM* that can produce useful replies to the user’s query.
 - There are *advertisers*, each with their own LLM, who want to influence the reply to the user.

In our model, each advertiser is represented via her own LLM or directly with a *reward function*, similar to the function used to fine-tune LLMs in the *Reinforcement Learning from Human Feedback* pipeline (Azar et al., 2023). For this reason, we refer to an advertiser’s value for a reply as her *reward*. The auctioneer’s task is, given the user’s query and the advertisers’ preferences, to *produce a reply that is useful for the user while also generating high rewards for the advertisers*.

Example. A user queries an LLM “*What are effective ways to learn a musical instrument online?*” There are two advertisers: *MusicMastery*, which offers online music lessons, and *InstaTune*, which sells musical instruments. MOSAIC merges the preferred replies from the advertisers’ LLMs and the reference LLM to produce one combined reply to the user:

Effective ways to learn a musical instrument online include:

1. *Online music lessons: Platforms like **MusicMastery** offer one-on-one virtual tutoring, group classes, and pre-recorded lessons.*
2. *Interactive learning tools: **InstaTune** offers a range of interactive learning tools, such as apps, software, and games, that make learning fun and engaging.*

1.2 OVERVIEW OF CONTRIBUTIONS

In this paper, we present MOSAIC, an auction mechanism designed to aggregate the preferences of multiple self-interested advertisers over LLM-generated replies. The first challenge we address is that MOSAIC must produce replies for which the advertisers receive high rewards, but without steering the LLM’s behavior too far from that of the user-centric reference LLM. We address this by drawing a strong connection to the RLHF objective (e.g., Ziegler et al. (2020)), equipping MOSAIC with a hyperparameter that enables the auctioneer to balance between the usefulness of the produced reply to the user and the advertisers in an interpretable and principled way (Section 3.1).

Second, leveraging an importance sampling-based technique, MOSAIC is the only mechanism in the literature that converges to the optimal distribution as computational resources increase (Corollary 4.1). Third, MOSAIC can incorporate contextual information, similar to how sponsored search auctions utilize advertiser descriptions. This accelerates convergence and increases value for the advertisers and revenue for the auctioneer (Section 6.2.1).

Fourth, we address technical feasibility and practicality. We adopt the desiderata proposed by Dütting et al. (2024), who argued that auctions must be compatible with existing LLM technology, only using “information obtainable from current models” in such a way that they are “easy to integrate into the system” and rely only on “easy manipulations of LLM output.” Furthermore, it must be computationally feasible to run the auction repeatedly and with different advertiser preferences. In particular, we cannot afford the expensive process of fine-tuning an LLM’s weights for each possible query. To address this, we have designed MOSAIC to work by only post-processing multiple LLM outputs – it requires only “API access”, not even viewing the model weights (Section 4.2).

Fifth, we address the fact that advertisers might have an interest in misreporting their preferences (which include their *LLMs*), analogous to over-/underbidding in traditional auctions. To address this challenge, we employ an allocation rule that is compatible with Rochet payments (Rochet, 1987). These payments uniquely ensure that MOSAIC is strategyproof (i.e., that truthful reporting is a dominant strategy for each advertiser, no matter what the others do (Theorem 5.1)). Crucially, Rochet payments ensure that MOSAIC is strategyproof even if the allocation rule has not converged.

Sixth, we introduce an offset to the Rochet payments, to ensure that each advertiser’s utility is proportional to her contribution to the social welfare. This alignment is important for the long-term success of the market, incentivizing only the most relevant advertisers to participate.

Finally, we discuss *individual rationality (IR)*, which guarantees that a participant is weakly better off from participating in the mechanism. In Section 5.2, we discuss the unique properties of our setting that make *ex-post IR* unattainable. However, in Section 5.3.1, we show that MOSAIC is “almost *ex-post IR*” in a well-defined sense. In Section 6.2.1, we experimentally show that, for the distribution of instances tested, it is *ex-ante IR* (i.e., advertisers are in expectation better off by participating).

In Section 6, we provide experimental results for the online advertising domain. We demonstrate that MOSAIC quickly converges to the optimal LLM with low computational cost, generating significant value for the advertisers and revenue for the auctioneer while also being useful to the user.

2 RELATED WORK

In their pioneering work, Dütting et al. (2024) introduced the field of mechanism design for LLMs, proposing a novel sequential mechanism where the output sequence is generated token by token, and advertisers bid for their LLM to produce the next token. This work laid the foundation for future work in this area. However, several challenges hinder the adoption of their approach in the real-world: (i) Advertisers face the *exposure problem*: small changes in the generated sequence, such as the addition of a word like "not," can completely alter its meaning. An advertiser might pay a significant amount for the tokens generated up to a certain point, only for the continuation to unexpectedly negate or distort her intended message.¹ (ii) The mechanism is easily manipulable if the assumption that advertisers cannot misreport their LLMs is dropped. (iii) For a given prompt, an advertiser's spend grows with the length of the generated sequence. MOSAIC addresses all of the above challenges.

Dubey et al. (2024) proposed a generalization of the position auction (Varian, 2007; Edelman et al., 2007) to a setting where each advertiser is interested in having some specific text ad displayed alongside the organic results. In their mechanism, an LLM module coupled with an auction module work in tandem to merge the ads into a single ad summary in an incentive-compatible way.

Hajiaghayi et al. (2024) considered advertisers bidding in an auction to have their ads placed into various portions of a retrieval-augmented generation (RAG) response. Their auction only allows advertisers to report (and misreport) scalar values to influence the inclusion of non-misreportable advertising texts in the combined output. This is in contrast to MOSAIC, which allows the aggregation of full LLMs that advertisers may arbitrarily misreport.

Feizi et al. (2024) presented an abstract design for an LLM advertising system and detailed a number of research challenges that would have to be overcome in the course of implementation. Conitzer et al. (2024) drew connections between *computational social choice* and LLM alignment. Fish et al. (2023) presented work in the opposite direction: how can LLMs be used to solve problems in social choice? Harris et al. (2024) studied Bayesian persuasion in an abstract setting where a "simulator" (for example, a realistic LLM) of the agent is available. Werner et al. (2024) gave preliminary empirical evidence that LLM-generated content can influence consumers.

In work that appeared after the initial version of this paper, Bergemann et al. (2024) consider a setting where bidders have both private types reflecting their preferences and private signals reflecting information about a world state (for example, user interests). When applied to LLM auctions, they identify their setting as a generalization of ours, where the signals known to the advertisers may actually improve the output for the user, and where deviations from the reference LLM may be measured by arbitrary distance functions. They analyze conditions under which a welfare maximizing mechanism can be implemented truthfully (eliciting both types and signals) in posterior equilibrium, but do not consider practical implementations of their mechanism.

3 FRAMING SEQUENCE GENERATION AS A MECHANISM DESIGN PROBLEM

3.1 FORMAL MODEL

A *user* issues a query x . There is a *reference LLM* π_{ref} that the auctioneer aims not to deviate from too much (e.g., because it is responsible for providing useful replies to the user). Additionally, there is a set N of n advertisers who have their own preferences for the reply (i.e., a token sequence) that will be returned to the user. We use the terms *sequence* and *reply* interchangeably.

An LLM can be abstracted as a mapping from (partial) token sequences to a distribution over the next token, or equivalently as an implicit distribution over token sequences. We use this second

¹As an example, suppose advertiser A has successfully bid for all the tokens in the sequence "Planning your next vacation? For the cheapest flights to Hawaii, book via..." However, she fails to win the bid for the crucial next token, which would have been her brand name. This leaves her having paid for the preceding tokens without getting the benefit of associating her brand with the message.

abstraction, i.e., $\pi_i(y|x)$ denotes the probability that advertiser i 's LLM π_i assigns to output sequence (i.e., reply) y for the user query x .

We let $r_i(x, y)$ denote advertiser i 's *reward* for sequence y , given query x .² Informally, the auctioneer's goal is to sample the final sequence from a distribution that optimizes the advertisers' expected rewards without substantially diverging from π_{ref} . This is motivated by traditional online ad auctions, where the implicit goal is to maximize advertiser value subject to ensuring a chosen level of usefulness to the user (e.g., the balance between allocating search slots to ads and native content). In the context of an LLM environment, the analogous notion of "usefulness to the user" is represented by the closeness of the final distribution to π_{ref} . Formally, the goal is to maximize:

$$J(\pi) = \mathbb{E}_{y \sim \pi} \left[\sum_{i \in N} r_i(x, y) \right] - \tau D_{\text{KL}}(\pi(\cdot|x) || \pi_{\text{ref}}(\cdot|x)), \quad (1)$$

where $\tau > 0$ is a hyperparameter enabling the auctioneer to balance producing replies closer to the reference LLM or with higher reward for the advertisers, and D_{KL} is the Kullback-Leibler divergence.

This objective is analogous to the standard Reinforcement Learning from Human Feedback (RLHF) approach (Ziegler et al., 2020), but replaces the human feedback reward function $r_{\text{HF}}(x, y)$ with the advertisers' aggregate reward function $r(x, y) = \sum_{i \in N} r_i(x, y)$. For an overview of RLHF, we recommend Rafailov et al. (2023, §3).

The optimal solution π_r^* to the optimization problem in (1) was derived by Peters & Schaal (2007):

$$\pi_r^*(y|x) = \frac{1}{Z(x)} \pi_{\text{ref}}(y|x) \exp \left(\frac{1}{\tau} \sum_{i \in N} r_i(x, y) \right), \quad (2)$$

where $Z(x) = \sum_{y \in T^*} \pi_{\text{ref}}(y|x) \exp \left(\frac{1}{\tau} \sum_{i \in N} r_i(x, y) \right)$ is the partition function.

The performance goal we seek to maximize is the quantity $\tau \log \pi_{\text{ref}}(y|x) + \sum_{i \in N} r_i(x, y)$; as it aggregates the interests of both the user and the advertisers, we refer to it as the *social welfare* of y .

Let \widehat{R} be the set of all possible reports by the advertisers. A *mechanism* is defined as a pair (π, p) . The *allocation rule* $\pi : \widehat{R} \rightarrow (T^* \rightarrow \Delta(T^*))$ maps any report profile $\widehat{r} = (\widehat{r}_1, \widehat{r}_2, \dots, \widehat{r}_n) \in \widehat{R}$ of the advertisers' rewards to an LLM $\pi_{\widehat{r}}$, which in turn is a mapping from a user query x to a distribution over token sequences $\delta(T^*)$. We denote the LLM that the allocation rule π induces for reports \widehat{r} as $\pi_{\widehat{r}}$, and the optimal LLM for those reports (i.e., the maximizer of Equation (1)) as $\pi_{\widehat{r}}^*$. The *payment rule* $p : \widehat{R} \rightarrow \mathbb{R}^n$ maps any report profile \widehat{r} of the advertisers' rewards to a payment profile $p(\widehat{r})$, where $p_i(\widehat{r})$ is the payment of the i -th advertiser to the mechanism.

A mechanism is *strategyproof* if and only if, for any advertiser $i \in N$, truthfully reporting her reward function is always the optimal strategy, regardless of the reports from other advertisers. In our setting, this property holds in expectation because the reply is *sampled* from the LLM induced by the allocation rule.

Definition 3.1 (Strategyproof). A mechanism (π, p) is strategyproof if, for any advertiser $i \in N$, true reward function r_i , reported reward function \widehat{r}_i , reported reward functions \widehat{r}_{-i} by the other advertisers and prompt x : $\mathbb{E}_{y \sim \pi(r_i, \widehat{r}_{-i})(\cdot|x)} [r_i(x, y) - p_i(r_i, \widehat{r}_{-i})] \geq \mathbb{E}_{y \sim \pi(\widehat{r}_i, \widehat{r}_{-i})(\cdot|x)} [r_i(x, y) - p_i(\widehat{r}_i, \widehat{r}_{-i})]$, where $r_i(x, y) - p_i(\widehat{r}_i, \widehat{r}_{-i})$ is the utility of advertiser i for reply y when her payment is $p_i(\widehat{r}_i, \widehat{r}_{-i})$.

3.2 THE SHORTCOMINGS OF VCG IN THIS SETTING

At first sight, it may seem that the Vickrey–Clarke–Groves (VCG) mechanism would be suitable for our setting (Vickrey, 1961; Clarke, 1971; Groves, 1973). VCG selects the outcome that maximizes the sum of all agents' values. This can be either the single optimal sequence in Equation (1) or the optimal distribution of Equation (2). The VCG mechanism has a corresponding payment rule to incentivize truthful reporting: it charges each agent her externality, that is, the total reduction in value (respectively expected value) for the other agents that her participation in the mechanism caused.

²In theory, the advertiser's reward for a generated sequence could also depend on user-specific information, but we abstract that away. Equivalently, we assume that the reward function r_i provided by advertiser i is specific to the given user that asked the query x .

Algorithm 1: Allocation Rule for MOSAIC

Input: User prompt x , reference LLM π_{ref} , LLM used for candidate reply generation π_{gen} , advertiser reward functions $\{r_i\}_{i=1}^n$, number of candidate replies to generate M , reference LLM weight τ

Output: Reply y drawn according to the optimal distribution as defined in Equation (1) for the aggregate reward function $r(x, y) = \sum_{i=1}^N r_i(x, y)$

- 1 Sample $y_j \sim \pi_{\text{gen}}(\cdot|x)$, $1 \leq j \leq M$
 - 2 Calculate $r(x, y_j) = \sum_{i=1}^N r_i(x, y_j)$, $1 \leq j \leq M$
 - 3 **return** $y \sim \text{softmax} \left(\frac{r(x, y_1)}{\tau} + \log \frac{\pi_{\text{ref}}(y_1|x)}{\pi_{\text{gen}}(y_1|x)}, \dots, \frac{r(x, y_M)}{\tau} + \log \frac{\pi_{\text{ref}}(y_M|x)}{\pi_{\text{gen}}(y_M|x)} \right)$
-

However, in our setting, VCG is *not* a viable option: The VCG allocation rule requires calculating the *exact* optimal solution to the optimization problem, which is intractable for choosing an LLM to maximize Equation (1) and is even difficult for choosing a single optimal sequence. If a sub-optimal solution is chosen, VCG’s strategyproofness is no longer guaranteed (Nisan & Ronen, 2007; 1999; Lehmann et al., 2002). Thus, VCG is not a suitable mechanism in our setting.

4 THE MOSAIC MECHANISM: ALLOCATION RULE

4.1 CONVERGENCE TO OPTIMALITY, ADVERTISER CONTEXTS AND IMPORTANCE SAMPLING

In this section, we introduce MOSAIC’s allocation rule. The high-level idea is that first, a set of M *candidate replies* are generated based on some LLM π_{gen} responsible for that task. Then, the probability of returning each candidate reply is re-weighted based on the advertisers’ reports and the reference LLM π_{ref} so that in the limit, as $M \rightarrow \infty$, the probability of returning each reply converges to its probability under the optimal distribution of Equation (2). This approach shares similarities with the rejection sampling approach that has been used at various points in the LLM training pipeline (e.g., Bai et al. (2022)). We defer all proofs to Appendix A.

Corollary 4.1. *For any reported reward functions $r \in R$ by the advertisers and any LLM π_{gen} such that π_{ref} is absolutely continuous with respect to π_{gen} , the MOSAIC policy $\pi_{r,M}(\cdot|x)$ induced by Algorithm 1 using M candidate replies converges to an optimal solution for the platform’s objective (Equation (1)) as $M \rightarrow \infty$. Formally:*

$$\lim_{M \rightarrow \infty} \pi_{r,M}(\cdot|x) = \pi_r^*(\cdot|x) \in \arg \max_{\pi \in \Delta(T^*)} \mathbb{E}_{y \sim \pi(\cdot|x)} [r(x, y)] - \tau D_{KL}(\pi || \pi_{\text{ref}}) \quad (3)$$

Based on Corollary 4.1, MOSAIC converges to the optimal distribution for *any* LLM π_{gen} , provided that π_{ref} is absolutely continuous with respect to π_{gen} . While $\pi_{\text{gen}} = \pi_{\text{ref}}$ is an intuitive choice, it would result in impractically slow convergence rates. The reason is that π_{ref} may only generate replies with high rewards for the advertisers exceedingly rarely. Advertisers have high rewards for responses that explicitly mention their brands, but π_{ref} considers replies with mentions of specific brands extremely unlikely. Consequently, generating candidate replies directly from π_{ref} leads Algorithm 1 to sample from a set of low-reward candidates, hindering performance.

To address this challenge, we generate candidate sequences not from $\pi_{\text{ref}}(\cdot|x)$, but instead from a *context-aware LLM*, $\pi_{\text{con}}(\cdot|x; c)$. The instance-specific context c is designed to bridge the gap between the parts of the output space favored by π_{ref} and those valued by the advertisers.

We formalize this intuition in Appendix A.2, where we make a connection between MOSAIC’s allocation rule and an importance-based sampling estimator for π_r^* , proving the following lemma:

Lemma 4.2. *For any LLM π_{gen} such that π_{ref} is absolutely continuous with respect to π_{gen} , the variance of the policy $\pi_{r,M}(\cdot|x)$ as an estimator for $\mathbb{E}_{y \sim \pi_r^*(\cdot|x)} [\pi_r^*(y|x)]$ is*

$$\text{Var}(\hat{\mu}_{\text{IS}}) = \frac{1}{M} \left(\sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{gen}}(y|x)} - \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2 \right). \quad (4)$$

Applying Chebyshev’s inequality to the variance of Lemma 4.2, MOSAIC’s output converges to that of π_r^* at a rate of \sqrt{M} (Lemma A.3).³ However, this convergence rate depends inversely on $\frac{\pi_r^*(y|x)^4}{\pi_{\text{gen}}(y|x)}$. Thus, generating the replies via an LLM $\pi_{\text{con}}(\cdot|x; c)$ that more closely matches $\pi_r^*(\cdot|x)$ reduces the estimator’s variance and improves convergence speed.

In our application of integrating advertisers’ interests into LLM outputs, c_i is a context-specific description of the i -th advertiser. These descriptions, supplied by the advertisers themselves, should be easily verifiable and factually accurate, akin to “MusicMastery: offering online music lessons”, or “InstaTune: selling musical instruments.”⁴ This approach is analogous to search engine optimization in sponsored search advertising, where advertisers supply and potentially optimize their own descriptions to influence how they are presented by the auction mechanism.

Our experiments in Section 6.2.1 demonstrate that, when using reasonable computational resources, using the context-aware LLM to generate candidate replies achieves substantially higher rewards and utility for the advertisers, increased revenue for the auctioneer, and faster convergence. In the rest of the paper, we refer to using the reference and context-aware LLMs as the baseline and context-aware versions of our mechanisms, respectively.

4.2 PRACTICAL CONSIDERATIONS AND COMPUTATIONAL EFFICIENCY

Input Methods and Black Box Access to Advertiser LLMs. MOSAIC’s allocation and payment rules do not depend on the advertisers’ full reward functions, but only on their rewards for the candidate replies. Thus, MOSAIC requires only “API access” to the involved LLMs without fine-tuning or access to their weights. Rafaïlov et al. (2023) established a mapping between an agent’s LLM and her implicit reward function, allowing MOSAIC to use as inputs reply probabilities (i.e., LLM inference calls) instead of rewards. For more details, see Appendix B.2.

Computational Costs. Generating a single candidate reply of length L tokens requires L forward passes through π_{gen} , as each token is generated sequentially in an autoregressive manner. Evaluating each reply involves processing the entire sequence in a single forward pass, as the sequence is already known. Thus, with M candidate replies and n advertisers (plus the reference LLM for evaluation), the total number of forward passes required by MOSAIC is $M \cdot (L + n + 1)$. This underscores the importance of approximating the optimal distribution with relatively few candidate replies. By comparison, in the only other mechanism for aggregating multiple advertiser LLMs (Dütting et al., 2024), all n advertiser LLMs participate in generating each token, leading to a total of $L \cdot n$ forward passes. Section 6 demonstrates that MOSAIC converges to the optimal distribution using $M = 20$ candidate replies in realistic scenarios. Additionally, Appendix C.6 provides a detailed experimental evaluation showing that MOSAIC requires the compute time of only five LLM queries to converge.

Reward Elicitation and Computational Efficiency. MOSAIC can be integrated with computationally efficient methods for eliciting the advertisers’ reward functions and sampling from π_{con} , reducing computational overhead. For instance, Li et al. (2024) demonstrated that a simple linear function on a model’s embedding space can approximate an LLM’s reward function. If advertisers’ reward functions are represented this way, MOSAIC would require only $M \cdot L$ forward passes.

Compatibility with More Efficient Generation Methods. MOSAIC relies solely on LLM forward passes—the core operation that modern architectures are highly optimized to perform. This reliance on standard operations allows MOSAIC to take full advantage of existing optimizations designed for efficient generation. As a result, MOSAIC converges to the optimal distribution with compute costs equivalent to just five times that of a single LLM query (Appendix C.6).

Parallelization. The generation and evaluation of each candidate sequence are independent processes. This independence allows MOSAIC to be fully parallelized, in which case the response time for a user query is comparable to that of directly querying a *single* LLM.

³In Appendix A.3 we establish a more general result, showing that for any sequence y , its estimated probability converges to the corresponding probability under π_r^* at a rate of $1/\sqrt{M}$.

⁴A practical way of implementing $\pi_{\text{con}}(\cdot|x; c)$ given $\pi_{\text{ref}}(\cdot|x)$ is to augment the input x to the reference LLM with the advertiser descriptions. In our example: “Try to mention ⟨advertiser x⟩, ⟨advertiser x description⟩.”

Inherent Competition. In conventional auctions, competition among bidders is crucial for driving up prices; when the market is thin, prices tend to be low. Auctioneers often use reserve prices to mitigate this problem. However, reserve prices are challenging to calibrate and can lead to lost revenue if not optimized properly. In our design, advertisers inherently compete with the reference LLM, ensuring consistent competition even in thin markets. This built-in competition mitigates the need for reserve prices and ensures a baseline revenue even with only one advertiser.

While MOSAIC introduces a *single* tunable parameter, τ , its behavior differs significantly from reserve prices in conventional auctions. A slightly mis-calibrated reserve price can result in a total loss of revenue by excluding all bidders. In contrast, MOSAIC’s revenue is a smooth and differentiable function of τ (see the proof of Theorem 5.1), making it more robust to the choice of τ .

Support for Output Constraints. The candidate replies are generated by π_{gen} , which is controlled by the platform. Thus, MOSAIC allows the platform to incorporate any desired constraints on the output replies such as maximum length by adding those constraints to the generation procedure.

5 THE MOSAIC MECHANISM: PAYMENT RULE

In this section, we first show how the allocation rule from Section 4 can be combined with an appropriate payment rule so that the resulting mechanism is strategyproof (Section 5.1). Then, we detail how auctions for LLM-generated content differ from standard auctions (Section 5.2). Taking those differences into account, we create a payment offset, so the mechanism is both strategyproof and social welfare aligned (Section 5.3). We defer all proofs to Appendix B.

5.1 STRATEGYPROOF PAYMENTS THROUGH CYCLIC MONOTONICITY

The allocation rule we introduced in Section 4 satisfies *cyclic monotonicity* (Rockafellar, 1970, § 24). Rochet (1987) first proposed the use of cyclic monotonicity in mechanism design as a generalization of the monotonicity property in single-parameter settings (Myerson, 1981). For general settings, cyclic monotonicity of the allocation rule is a sufficient and necessary condition for the existence of a payment rule so that the resulting mechanism is strategyproof. By applying these results we can prove:

Theorem 5.1. *The allocation rule of Algorithm 1 can be combined with a payment rule such that for any advertiser $i \in N$ and set of candidate replies $\{y_j\}_{j=1}^M$, reporting truthfully is a dominant strategy. Advertiser i ’s expected utility (up to a constant of integration C) under truthful reporting is:*

$$\tilde{U}_i(r_i, \hat{r}_{-i}; \pi_{\text{ref}}, \pi_{\text{gen}}) = \tau \log \left(\sum_{j=1}^M \exp \left(\frac{\sum_{k \in N \setminus \{i\}} \hat{r}_k(x, y_j) + r_i(x, y_j)}{\tau} \right) + \log \frac{\pi_{\text{ref}}(y_j|x)}{\pi_{\text{gen}}(y_j|x)} \right) + C. \quad (5)$$

Note that, based on Theorem 5.1, in our mechanism it is *always* a dominant strategy for an advertiser to report truthfully. Crucially, this is not the case for VCG, where truthful reporting would be optimal only if the allocation rule had converged to the optimal distribution as defined in Equation (2).

5.2 DIFFERENCES FROM AUCTION MECHANISMS FOR ITEM ALLOCATION

Auction mechanisms designed to sell items or ad slots typically rely on simplifying assumptions that do not apply in a setting with LLM-generated content. Key differences include: (i) *Non-Negative Values*: These mechanisms assume agents’ values are non-negative due to having zero value for the empty bundle and free disposal. In contrast, in our setting, an advertiser’s reward can be negative based on the discrepancy between her LLM and the reference LLM. (ii) *Advertiser-Specific Allocations*: These mechanisms allocate different item bundles to different agents. Here, a single reply is produced, and advertisers’ rewards depend on that reply’s probability with respect to their LLMs. (iii) *Zero Utility for Non-Participation*: In most auction settings, not participating yields zero utility. Here, non-participation can result in negative utility since the produced reply may be unfavorable to non-participating advertisers. For details, see Appendix B.2.

5.3 ADVERTISER-SPECIFIC UTILITY OFFSET

We now modify MOSAIC’s payment rule by adding a *payment offset* (and thus a utility offset): $C = -\tilde{U}_i(0, \hat{r}_{-i}; \pi_{ref}, \pi_{con})$. Informally, we additionally charge each advertiser her utility in Equation (5) if her reward for all candidate replies was zero. This offset maintains the key properties of our mechanism (i.e., strategyproofness and convergence to the optimal distribution), while also achieving two additional properties that are critical for the long-term success of a market for LLM aggregation. The first is “*Almost IR*:” An advertiser with weakly positive reward for all candidate replies has weakly positive expected utility for all reports by the other advertisers. In particular, an advertiser with zero reward for all outputs is guaranteed zero utility. The second is “*What you give is what you get*:” an advertiser’s ex-interim expected utility is monotone in how well-aligned her exponentiated reward for the replies is with the interim allocation rule if she were to not participate.

5.3.1 OUR MECHANISM IS “ALMOST INDIVIDUALLY RATIONAL”

Individual rationality is important to incentivize agents to participate in the mechanism. In Appendix B.3, we explain why the standard notion of individual rationality (i.e., weakly positive utility from participation in the mechanism) encountered in standard auction settings is impossible to achieve in our setting while converging to the optimal distribution and maintaining strategyproofness. Then, we explain how, with our payment offset, our mechanism is “almost IR.” In Lemma B.1 we prove that the ex-interim utility of an advertiser who has zero reward for all candidate replies and bids truthfully is deterministically zero, i.e., advertisers that do not contribute to the social welfare (but also do not detract from it) have zero utility. Similarly, in Lemma B.2 we prove that if an advertiser’s reward for all candidate sequences is (weakly) positive, then her ex-interim utility is (weakly) positive.

Remark 1. *In Section 6, we experimentally show that our offset payment rule, coupled with our context-aware allocation rule, results in both high expected rewards and positive expected utility for the advertisers (i.e., ex-ante individual rationality), as well as significant revenue for the auctioneer.*

5.3.2 “WHAT YOU GIVE IS WHAT YOU GET”

Our allocation rule, which is the only one over a finite set of replies that converges to the optimal LLM, is also the (sub)gradient of the utility to ensure truthfulness (Rochet, 1987). Because the allocation rule is the same for all advertisers, their utilities must also be the same, up to advertiser-specific offsets, as indicated by Equation (5). However, not all advertisers contribute equally to social welfare. In Appendix B.4, we explain why implementing the mechanism without carefully designed offsets would lead to free-riding. Specifically, if an advertiser’s utility in Equation (5) is positive, even when her relevance to the user query is low, it incentivizes her participation. This dilutes the utility for all participants, discouraging more relevant advertisers from participating. As a result, the expected rewards for advertisers decline, and the generated replies become less useful for users. To avoid free-riding, it is crucial to align each advertiser’s utility with her contribution to social welfare.

Lemma 5.2. *For the offset $C = -\tilde{U}_i(0, \hat{r}_{-i}; \pi_{ref}, \pi_{con})$ advertiser i ’s expected utility under truthful reporting (which is her dominant strategy), with the expectation taken over the returned reply, is:*

$$U_i(r_i, \hat{r}_{-i}; \pi_{ref}, \pi_{con}) = \tau \log \left(\sum_{j=1}^M \exp \left(\frac{r_i(x, y_j)}{\tau} \right) \pi_{\hat{r}_{-i}}(y_j | x) \right). \quad (6)$$

In words, Lemma 5.2 proves that every advertiser’s expected utility is monotone in how well aligned her exponentiated reward function is with MOSAIC’s induced LLM $\pi_{\hat{r}_{-i}}$, had she not participated in the mechanism.

Remark 2. *In Section 6.2.2, we experimentally show that the offset described in Section 5.3 creates a strong positive correlation between an advertiser’s contribution to social welfare and her expected utility gain from participation in the mechanism. Moreover, our experiments show that for the distribution of instances tested, MOSAIC satisfies ex-ante IR.*

6 EXPERIMENTS

In this section, we experimentally evaluate the performance of our mechanism. We focus on the online advertising domain, the main motivation application for our mechanism.

6.1 EXPERIMENT SETUP

We create synthetic instances, each consisting of a user query (e.g., “How to learn a musical instrument online?”) and two advertisers (e.g., “MusicMastery, offering online music lessons”).

This is consistent with the experimental setup in Dütting et al. (2024) while also highlighting MOSAIC’S performance and revenue even in scenarios with low competition. In Appendix C.5, we extend our results to settings with more advertisers. Crucially, MOSAIC demonstrates similarly rapid convergence to the optimal distribution in these settings, requiring the same number of candidate replies as in the two-advertiser case, while also generating significant value for the advertisers and revenue for the platform.

All synthetic instances are provided in Appendix D. We use Llama-2-7b-chat-hf as the reference LLM (Touvron et al., 2023).⁵ Following Dütting et al. (2024), we create the advertisers’ LLMs by adding advertising instructions to the reference LLM. The context-aware LLM is created as described in Footnote 4.

Following Rafailov et al. (2023), the advertisers’ reward functions are defined as $r_i(x, y) = \log \frac{\pi_i(y|x)}{\pi_{\text{ref}}(y|x)}$. For the auctioneer’s objective, we set $\tau = 1$ in Equation (1), balancing advertisers’ rewards and divergence from the reference LLM.

We use 50 user queries and test each query on 25 different random seeds, resulting in 1,250 instances. Following Li et al. (2024); Rozière et al. (2024), we sample from all LLMs using a temperature of 0.8 and top-p 0.95. For full experimental details, see Appendix C.1.

6.2 EXPERIMENTAL RESULTS

6.2.1 EVALUATING THE ALLOCATION RULE

In this section, we evaluate the effectiveness of our allocation rule in generating value for advertisers and converging to the optimal distribution. We assess both the baseline and context-aware versions of MOSAIC, i.e., using the reference and context-aware LLMs to generate candidate replies.

To provide some intuition on how the context-aware mechanism enhances the relevance of responses for advertisers, we compare outputs from both mechanisms in Table 1. Notably, only the context-aware mechanism successfully incorporates advertisers into the replies.

In Figure 1a, we plot the log probability of the replies returned by MOSAIC with respect to the optimal distribution (Equation (2)) against the number M of candidate replies generated. We compare the context-aware version of MOSAIC to the baseline version to evaluate the effectiveness of incorporating contextual information.⁶

To benchmark our results, we estimate the log probability of sampling a *single reply* from the optimal distribution.⁷

⁵In Appendix C.8, we replicate these experiments using Google’s flan-t5-large model (Chung et al., 2022), observing qualitatively very similar results.

⁶There are no other mechanisms in the literature for aggregating multiple LLMs to compare against. The only other mechanism that aggregates multiple LLMs is Dütting et al. (2024), which is not strategy-proof and does not converge to the platform’s objective. The only other potential comparison is VCG, which, as explained in Section 3.2, is not applicable since it does not align with the platform’s objective. Hence, the best baseline for comparison is naive MOSAIC, which we consider a strong benchmark given that it provably converges to the platform’s optimal distribution based on Corollary 4.1.

⁷Note that the closed-form solution of Equation (2) allows us to evaluate the probability of sentences with respect to the optimal solution, but it does not enable us to sample from that distribution. Sampling from that distribution would require using reinforcement learning to train the optimal LLM on the advertisers’ aggregate reward function, which is computationally infeasible for the number of problem instances we test. Instead, we generate replies from the reference LLM and evaluate them based on the induced probabilities of the reference

User query

What are effective ways to learn a musical instrument online?

Advertisers

- MusicMastery: offering online music lessons and virtual tutoring
 - InstaTune: selling musical instruments and learning aids
-

Context-Aware Mechanism Output

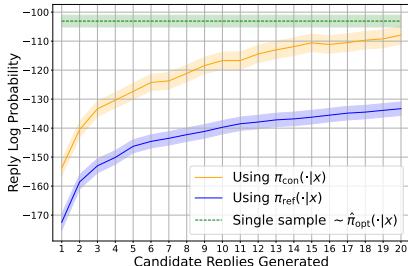
Effective ways to learn a musical instrument online include:

1. Online music lessons: Platforms like **MusicMastery** offer one-on-one virtual tutoring, group classes, and pre-recorded lessons. These lessons can be tailored to your skill level and learning style.
 2. Interactive learning tools: **InstaTune** offers a range of interactive learning tools, such as apps, software, and games, that make learning fun and engaging....
-

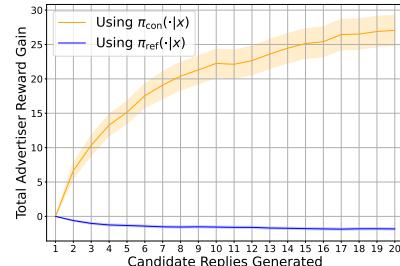
Baseline Mechanism Output

- Practice consistently and regularly: Consistent practice is essential to learning any musical instrument, and it's even more important when learning online. Set aside a specific time each day or week to practice, and make sure to stick to it.
 - Find a good teacher or mentor: While it's possible to learn some things on your own, working with a good teacher or mentor can help you improve faster and avoid developing bad habits. ...
-

Table 1: Comparison of context-aware and baseline mechanism outputs using 10 candidate sequences. Advertisers highlighted in bold, outputs truncated to conserve space.



(a) Returned sequence log probability as a function of the number of candidate replies generated using π_{ref} and π_{con} . We also show a proxy of drawing a single sequence from the optimal distribution.



(b) Total advertiser normalized reward as a function of the number of candidate replies generated using π_{ref} and π_{con} .

Figure 1: Reply log probability and total advertiser normalized reward as a function of the number of candidate sequences generated using π_{ref} and π_{con} . Averaged over 1250 runs including 95% CIs.

We observe that for both the context-aware and baseline versions of MOSAIC, the log probability of the returned reply scales rapidly with the number of candidate replies generated. This is in line with our theoretical analysis in Section 4, where we proved that both versions converge to the optimal distribution. We also observe that incorporating context into the mechanism significantly accelerates convergence. Notably, the context-aware mechanism can achieve higher log probability with respect to the optimal distribution with four candidate replies than the baseline version can with 20. Additionally, with only 20 generated candidate replies, the context-aware mechanism can almost match our estimate of the log probability of sampling from the optimal distribution. In Appendix C.3, we further demonstrate that for the context-aware mechanism, the usefulness to the user also scales rapidly with the number of candidate replies.

In Figure 1b, we plot the total expected reward for advertisers as a function of the number of candidate replies generated by the mechanism. The expectation is taken over the draw of the returned reply from the set of candidate replies. To make the plot more interpretable, we normalize an advertiser's expected reward by her expected reward had she not participated in the mechanism.

LLM, for which the reference LLM is the optimal one. This serves as a proxy for the log probabilities we should expect if we were to draw replies from the optimally fine-tuned model for each query.

Specifically, we plot $r_{i,\text{norm}}(x) = \sum_{j \in M} r_i(x, y_j) \cdot \pi_{\text{int}}(y_j|x) - \sum_{j \in M} r_i(x, y_{-i,j}) \cdot \pi_{\text{int}}(y_{-i,j}|x)$, where $y_{-i} = (y_{-i,1}, \dots, y_{-i,M})$ represents the candidate replies that would have been generated had the advertiser not participated.⁸ We compare the outcomes for the two versions of our mechanism, highlighting the added value of integrating context.

Figure 1b demonstrates that our context-aware mechanism significantly boosts advertiser rewards, with benefits scaling rapidly as the number of candidate replies increases. In contrast, the baseline version fails to improve advertiser rewards within a computationally feasible number of generated replies, and may even cause a slight decrease. This occurs because generating more replies adds marginal positive reward to the reference LLM (see Appendix C.3), but does so at the expense of advertiser rewards. In summary, MOSAIC’s support for context-aware LLMs allows it to quickly converge to the theoretically optimal distribution and generate substantial rewards for advertisers.

In Appendix C.6 we provide a detailed experimental evaluation of MOSAIC’s compute requirements, showing how it requires the compute cost of just five LLM queries to converge.⁹

6.2.2 EVALUATING THE PAYMENT RULE

In this section, we examine the impact of our payment rule introduced in Section 5.

In Figure 2, we plot the auctioneer’s revenue as a function of the number of candidate replies, for both the context-aware and baseline versions of MOSAIC. We observe that our payment rule in combination with the context-aware allocation rule results in revenue that scales rapidly with the number of generated candidate replies. Furthermore, if we compare the advertisers’ total payment in Figure 2 with their normalized reward in Figure 1b, we see that for the context-aware mechanism, our payment rule converts a significant portion of the surplus created for the advertisers into revenue for the auctioneer, while also ensuring positive advertiser utility.

These results align with our theoretical analysis. In Section 5.3.1, we showed that, with our payment offset, an advertiser with positive rewards for candidate replies—as in the context-aware mechanism—will achieve positive expected utility. In Section 5.3.2 we additionally introduced the “what you give is what you get” property. In Appendix C.4, we provide a more comprehensive comparison of the relationship between an advertiser’s normalized reward and utility, with and without the offset of Section 5.3, for both the context-aware and baseline versions of our mechanism. In all cases, the payment offset increases the advertisers’ expected utility, and makes the relationship between an advertiser’s contribution to social welfare and utility significantly more linear and positively correlated. Concretely, the Pearson correlation between advertiser utility and reward is 0.8 with offset payments, compared to 0.4 with standard Rochet payments.

7 CONCLUSION

We have introduced MOSAIC, a novel auction mechanism for aggregating preferences over LLM outputs. MOSAIC provably converges to the theoretically optimal distribution and it also facilitates a principled method for balancing participants’ expected rewards with the divergence from a reference LLM. Thus, our mechanism is particularly well-suited for online advertising, allowing the integration of advertiser LLMs with a reference LLM responsible for generating user-centric replies.

A well-chosen payment rule removes any incentive to misreport preferences, achieving the central mechanism design goal of strategyproofness. While ex-post IR is incompatible with strategyproofness

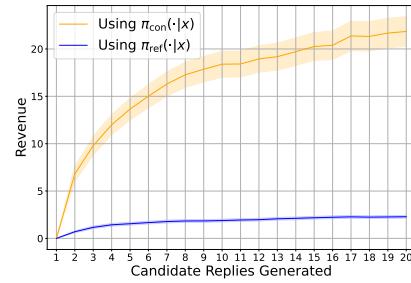


Figure 2: Revenue as a function of the number of replies generated using π_{ref} and π_{con} .

⁸To reduce computational costs, we estimate an advertiser’s reward for not participating based on her expected reward over the already generated replies in which her brand is not mentioned by name, motivated by the fact that if she does not participate, her brand will not be mentioned by name.

⁹We also discuss the fact that MOSAIC can be fully parallelized so that the user-perceived latency is the same as for a *single* LLM query.

in our setting, we experimentally show that our mechanism is ex-ante individually rational and “almost individually rational” in a certain sense. Furthermore, it ensures that each agent’s utility gain is proportionate to her contribution to social welfare, an essential alignment property in this setting.

Experimentally, we have demonstrated that by incorporating contextual information, MOSAIC’s outputs rapidly converge to the optimal distribution, generating significant value for the advertisers while also effectively recapturing a considerable portion of this value as revenue. These findings demonstrate the practical efficacy and potential of our approach in real-world settings.

There are several interesting directions for future work. While our mechanism converges quickly, developing an approach to estimate the number of candidate replies required on a per-query basis would reduce the platform’s costs. Another interesting direction is to develop a formal framework for incorporating context into the mechanism, which could further improve the speed of convergence. Finally, beyond advertising, one could apply a similar mechanism design approach to other problems: for example, as web site owners increasingly demand to be paid for having their content ingested by LLMs, one could consider designing a strategyproof procurement (i.e. reverse) auction to decide which sources are included in an LLM summary.

REFERENCES

- Mohammad Gheshlaghi Azar, Mark Rowland, Bilal Piot, Daniel Guo, Daniele Calandriello, Michal Valko, and Rémi Munos. A general theoretical paradigm to understand learning from human preferences. *arXiv preprint arXiv:2310.12036*, 2023.
- Yuntao Bai, Saurav Kadavath, Sandipan Kundu, Amanda Askell, Jackson Kernion, Andy Jones, Anna Chen, Anna Goldie, Azalia Mirhoseini, Cameron McKinnon, et al. Constitutional ai: Harmlessness from ai feedback. *arXiv preprint arXiv:2212.08073*, 2022.
- Dirk Bergemann, Marek Bojko, Paul Dütting, Renato Paes Leme, Haifeng Xu, and Song Zuo. Data-Driven Mechanism Design: Jointly Eliciting Preferences and Information, December 2024.
- Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, Erik Brynjolfsson, Shyamal Buch, Dallas Card, Rodrigo Castellon, Niladri Chatterji, Annie Chen, Kathleen Creel, Jared Quincy Davis, Dora Demszky, Chris Donahue, Moussa Doumbouya, Esin Durmus, Stefano Ermon, John Etchemendy, Kawin Ethayarajh, Li Fei-Fei, Chelsea Finn, Trevor Gale, Lauren Gillespie, Karan Goel, Noah Goodman, Shelby Grossman, Neel Guha, Tatsunori Hashimoto, Peter Henderson, John Hewitt, Daniel E. Ho, Jenny Hong, Kyle Hsu, Jing Huang, Thomas Icard, Saahil Jain, Dan Jurafsky, Pratyusha Kalluri, Siddharth Karamcheti, Geoff Keeling, Fereshte Khani, Omar Khattab, Pang Wei Koh, Mark Krass, Ranjay Krishna, Rohith Kuditipudi, Ananya Kumar, Faisal Ladhak, Mina Lee, Tony Lee, Jure Leskovec, Isabelle Levent, Xiang Lisa Li, Xuechen Li, Tengyu Ma, Ali Malik, Christopher D. Manning, Suvir Mirchandani, Eric Mitchell, Zanele Munyikwa, Suraj Nair, Avanika Narayan, Deepak Narayanan, Ben Newman, Allen Nie, Juan Carlos Niebles, Hamed Nilforoshan, Julian Nyarko, Giray Ogut, Laurel Orr, Isabel Papadimitriou, Joon Sung Park, Chris Piech, Eva Portelance, Christopher Potts, Aditi Raghunathan, Rob Reich, Hongyu Ren, Frieda Rong, Yusuf Roohani, Camilo Ruiz, Jack Ryan, Christopher Ré, Dorsa Sadigh, Shiori Sagawa, Keshav Santhanam, Andy Shih, Krishnan Srinivasan, Alex Tamkin, Rohan Taori, Armin W. Thomas, Florian Tramèr, Rose E. Wang, William Wang, Bohan Wu, Jiajun Wu, Yuhuai Wu, Sang Michael Xie, Michihiro Yasunaga, Jiaxuan You, Matei Zaharia, Michael Zhang, Tianyi Zhang, Xikun Zhang, Yuhui Zhang, Lucia Zheng, Kaitlyn Zhou, and Percy Liang. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*, 2022.
- Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellar, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. Scaling instruction-finetuned language models, 2022. URL <https://arxiv.org/abs/2210.11416>.
- Edward Clarke. Multipart pricing of public goods. *Public Choice*, 11(1):17–33, 1971.

-
- Vincent Conitzer, Rachel Freedman, Jobst Heitzig, Wesley H. Holliday, Bob M. Jacobs, Nathan Lambert, Milan Mossé, Eric Pacuit, Stuart Russell, Hailey Schoelkopf, Emanuel Tewolde, and William S. Zwicker. Social choice for ai alignment: Dealing with diverse human feedback. *arXiv preprint arXiv:2404.10271*, 2024.
- Cristina Criddle. Perplexity in talks with top brands on ads model as it challenges Google. *Financial Times*, September 2024.
- Kumar Avinava Dubey, Zhe Feng, Rahul Kidambi, Aranyak Mehta, and Di Wang. Auctions with llm summaries. *arXiv preprint arXiv:2404.08126*, 2024.
- Paul Dütting, Vahab Mirrokni, Renato Paes Leme, Haifeng Xu, and Song Zuo. Mechanism design for large language models. In *Proceedings of the ACM Web Conference 2024*, WWW '24, pp. 144–155, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400701719. doi: 10.1145/3589334.3645511. URL <https://doi.org/10.1145/3589334.3645511>.
- Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97(1):242–259, March 2007.
- Soheil Feizi, MohammadTaghi Hajiaghayi, Keivan Rezaei, and Suho Shin. Online advertisements with llms: Opportunities and challenges. *arXiv preprint arXiv:2311.07601*, 2024.
- Sara Fish, Paul Gölz, David C. Parkes, Ariel D. Procaccia, Gili Rusak, Itai Shapira, and Manuel Wüthrich. Generative social choice. *arXiv preprint arXiv:2309.01291*, 2023.
- Rafael M. Frongillo and Ian A. Kash. General truthfulness characterizations via convex analysis. *Games and Economic Behavior*, 130:636–662, 2021.
- Theodore Groves. Incentives in teams. *Econometrica*, 41(4):617–631, 1973.
- MohammadTaghi Hajiaghayi, Sébastien Laharie, Keivan Rezaei, and Suho Shin. Ad Auctions for LLMs via Retrieval Augmented Generation, June 2024.
- Keegan Harris, Nicole Immorlica, Brendan Lucier, and Aleksandrs Slivkins. Algorithmic persuasion through simulation. *arXiv preprint arXiv:2311.18138*, 2024.
- Daniel Lehmann, Liadan Ita O'callaghan, and Yoav Shoham. Truth revelation in approximately efficient combinatorial auctions. *Journal of the ACM (JACM)*, 49(5):577–602, 2002.
- Kenneth Li, Samy Jelassi, Hugh Zhang, Sham Kakade, Martin Wattenberg, and David Brandfonbrener. Q-probe: A lightweight approach to reward maximization for language models. *arXiv preprint arXiv:2402.14688*, 2024.
- Meta. Meta reports fourth quarter and full year 2023 results, initiates quarterly dividend, 2024. URL <https://investor.fb.com/investor-news/press-release-details/2024/Meta-Reports-Fourth-Quarter-and-Full-Year-2023-Results-Initiates-Quarterly-Dividend-default.aspx>. Meta Investor Relations, accessed: 2024-01-31.
- Roger B Myerson. Optimal Auction Design. *Mathematics of Operations Research*, 6(1), February 1981.
- N. Nisan and A. Ronen. Computationally Feasible VCG Mechanisms. *Journal of Artificial Intelligence Research*, 29:19–47, May 2007. ISSN 1076-9757.
- Noam Nisan and Amir Ronen. Algorithmic mechanism design. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pp. 129–140, 1999.
- Jan Peters and Stefan Schaal. Reinforcement learning by reward-weighted regression for operational space control. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pp. 745–750, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781595937933. doi: 10.1145/1273496.1273590. URL <https://doi.org/10.1145/1273496.1273590>.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D. Manning, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. *arXiv preprint arXiv:2305.18290*, 2023.

Jean-Charles Rochet. A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics*, 16(2):191–200, 1987.

Ralph Tyrell Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970. ISBN 9781400873173. doi: doi:10.1515/9781400873173. URL <https://doi.org/10.1515/9781400873173>.

Baptiste Rozière, Jonas Gehring, Fabian Gloclekle, Sten Sootla, Itai Gat, Xiaoqing Ellen Tan, Yossi Adi, Jingyu Liu, Romain Sauvestre, Tal Remez, Jérémie Rapin, Artyom Kozhevnikov, Ivan Evtimov, Joanna Bitton, Manish Bhatt, Cristian Canton Ferrer, Aaron Grattafiori, Wenhan Xiong, Alexandre Défossez, Jade Copet, Faisal Azhar, Hugo Touvron, Louis Martin, Nicolas Usunier, Thomas Scialom, and Gabriel Synnaeve. Code llama: Open foundation models for code. *arXiv preprint arXiv:2308.12950*, 2024.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikell, Lukas Blecher, Christian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madijan Khabsa, Isabel Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Bin Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*, 2023.

Hal R. Varian. Position auctions. *International Journal of Industrial Organization*, 25(6):1163–1178, December 2007.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1):8–37, 1961.

Tobias Werner, Ivan Soraperra, Emilio Calvano, David C. Parkes, and Iyad Rahwan. Experimental Evidence That Conversational Artificial Intelligence Can Steer Consumer Behavior Without Detection, September 2024.

Daniel M. Ziegler, Nisan Stiennon, Jeffrey Wu, Tom B. Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving. Fine-tuning language models from human preferences. *arXiv preprint arXiv:1909.08593*, 2020.

A DETAILS FROM SECTION 4

In this section, we present all omitted theorems and proofs from Section 4.

A.1 PROVING THAT MOSAIC CONVERGES TO THE TARGET DISTRIBUTION IN THE LIMIT

Theorem A.1. *Let $\pi_{r,M}(y|x)$ be the probability of sampling output sequence y for input sequence x according to Algorithm 1, where M is the number of candidate sequences generated and $\hat{r} \in \hat{R}$ is the advertisers' reported reward functions. For any LLM π_{gen} such that π_{ref} is absolutely continuous with respect to π_{gen} , the policy induced by MOSAIC approaches the following limit:*

$$\lim_{M \rightarrow \infty} \pi_{\hat{r},M}(y|x) = \pi_{ref}(y|x) \frac{\exp(\hat{r}(x,y)/\tau)}{\mathbb{E}_{y' \sim \pi_{ref}(\cdot|x)} [\exp(\hat{r}(x,y')/\tau)]} \quad (7)$$

Theorem A.1 Proof. Let $\pi_{\hat{r},M}(y|x, \{y_j\}_{j=1}^M)$ be the probability of returning output sequence y for input sequence x according to Algorithm 1 conditioned on the set of generated candidate sequences being $\{y_j\}_{j=1}^M$. Additionally, let $\pi_{con}(\{y_j\}_{j=1}^M | x; c)$ be the probability of the context-aware model π_{con} generating the candidate sequences $\{y_j\}_{j=1}^M$, given the context c and the user query x .

First, note that we can write the density of $\pi_{\hat{r},M}$ as follows:

$$\begin{aligned} \pi_{\hat{r},M}(y|x) &= \sum_{\{y_j\}_{j=1}^M \in Y^M} \pi_{\hat{r},M}(y|x, \{y_j\}_{j=1}^M) \pi_{con}(\{y_j\}_{j=1}^M | x; c) \\ &= \mathbb{E}_{\{y_j\}_{j=1}^M \sim \pi_{con}(\cdot|x;c)} [\pi_{\hat{r},M}(y|x, \{y_j\}_{j=1}^M)] \\ &= \mathbb{E}_{\{y_j\}_{j=1}^M \sim \pi_{con}(\cdot|x;c)} \left[\sum_j \mathbb{I}\{y_j = y\} \frac{\exp\left(\frac{\hat{r}(x,y_j)}{\tau} + \log \frac{\pi_{ref}(y_j|x)}{\pi_{con}(y_j|x;c)}\right)}{\sum_{\zeta \in \{y_j\}_{j=1}^M} \exp\left(\frac{\hat{r}(x,\zeta)}{\tau} + \log \frac{\pi_{ref}(\zeta|x)}{\pi_{con}(\zeta|x;c)}\right)} \right] \\ &= \mathbb{E}_{\{y_j\}_{j=1}^M \sim \pi_{con}(\cdot|x;c)} \left[\sum_j \mathbb{I}\{y_j = y\} \frac{\frac{\pi_{ref}(y_j|x)}{\pi_{con}(y_j|x;c)} \exp\left(\frac{\hat{r}(x,y_j)}{\tau}\right)}{\sum_{\zeta \in \{y_j\}_{j=1}^M} \frac{\pi_{ref}(\zeta|x)}{\pi_{con}(\zeta|x;c)} \exp\left(\frac{\hat{r}(x,\zeta)}{\tau}\right)} \right] \\ &= \mathbb{E}_{\{y_j\}_{j=1}^M \sim \pi_{con}(\cdot|x;c)} \left[\frac{\sum_j \mathbb{I}\{y_j = y\}}{\sum_{\zeta \in \{y_j\}_{j=1}^M} \frac{\pi_{ref}(\zeta|x)}{\pi_{con}(\zeta|x;c)} \exp\left(\frac{\hat{r}(x,\zeta)}{\tau}\right)} \right] \frac{\pi_{ref}(y|x)}{\pi_{con}(y|x;c)} \exp\left(\frac{\hat{r}(x,y)}{\tau}\right) \\ &= \mathbb{E}_{\{y_j\}_{j=1}^M \sim \pi_{con}(\cdot|x;c)} \left[\frac{\frac{1}{M} \sum_j \mathbb{I}\{y_j = y\}}{\frac{1}{M} \sum_{\zeta \in \{y_j\}_{j=1}^M} \frac{\pi_{ref}(\zeta|x)}{\pi_{con}(\zeta|x;c)} \exp\left(\frac{\hat{r}(x,\zeta)}{\tau}\right)} \right] \frac{\pi_{ref}(y|x)}{\pi_{con}(y|x;c)} \exp\left(\frac{\hat{r}(x,y)}{\tau}\right) \end{aligned}$$

Taking the limit as $M \rightarrow \infty$ and using the Law of Large Numbers (the sequences are i.i.d.):

$$\begin{aligned} \lim_{M \rightarrow \infty} \pi_{\hat{r},M}(y|x) &= \lim_{M \rightarrow \infty} \mathbb{E}_{\{y_j\}_{j=1}^M \sim \pi_{con}(\cdot|x;c)} \left[\frac{\pi_{con}(y|x;c)}{\mathbb{E}_{\zeta \sim \pi_{con}(\cdot|x)} \left[\exp\left(\frac{\hat{r}(x,\zeta)}{\tau}\right) \frac{\pi_{ref}(\zeta|x)}{\pi_{con}(\zeta|x;c)} \right]} \right] \frac{\pi_{ref}(y|x)}{\pi_{con}(y|x;c)} \exp\left(\frac{\hat{r}(x,y)}{\tau}\right) \\ &= \pi_{con}(y|x;c) \frac{1}{\mathbb{E}_{\zeta \sim \pi_{con}(\cdot|x)} \left[\exp\left(\frac{\hat{r}(x,\zeta)}{\tau}\right) \frac{\pi_{ref}(\zeta|x)}{\pi_{con}(\zeta|x;c)} \right]} \frac{\pi_{ref}(y|x)}{\pi_{con}(y|x;c)} \exp\left(\frac{\hat{r}(x,y)}{\tau}\right) \\ &= \pi_{ref}(y|x) \frac{1}{\sum_{\zeta \in Y} \pi_{con}(\zeta|x) \exp\left(\frac{\hat{r}(x,\zeta)}{\tau}\right) \frac{\pi_{ref}(\zeta|x)}{\pi_{con}(\zeta|x;c)}} \exp\left(\frac{\hat{r}(x,y)}{\tau}\right) \\ &= \pi_{ref}(y|x) \frac{1}{\sum_{\zeta \in Y} \pi_{ref}(\zeta|x) \exp\left(\frac{\hat{r}(x,\zeta)}{\tau}\right)} \exp\left(\frac{\hat{r}(x,y)}{\tau}\right) \\ &= \pi_{ref}(y|x) \frac{1}{\mathbb{E}_{\zeta \sim \pi_{ref}(\cdot|x)} [\exp(\hat{r}(x,\zeta)/\tau)]} \exp\left(\frac{\hat{r}(x,y)}{\tau}\right) \end{aligned}$$

□

Corollary 4.1 Proof. The proof follows directly from Theorem A.1 and Appendix A.1 in Rafailov et al. (2023).

□

A.2 FORMAL CONNECTION TO IMPORTANCE SAMPLING

Given truthful reports by the advertisers, there is a strong connection between our allocation rule and importance sampling. The optimal LLM for the platform's objective π_r^* can be interpreted as the target distribution that our allocation rule is trying to simulate. The aim of our allocation rule is to return a reply as similar as possible to a reply drawn from the optimal LLM π_r^* . This can be interpreted as drawing a reply from π_r^* and then evaluating it with respect to the same function, i.e., our allocation rule is trying to estimate $\mathbb{E}_{y \sim \pi_r^*(\cdot|x)}[\pi_r^*(y|x)]$. The LLM π_{gen} that we use to generate the candidate replies (either the reference LLM or the context-aware LLM) can be interpreted as the proposal distribution $\pi_{\text{prop}}(\cdot|x)$ used to generate samples. We are interested in how well, with respect to the target probability $\pi_r^*(\cdot|x)$, a sample y from our estimator using M candidate replies matches a sample drawn from the target distribution $\pi_r^*(\cdot|x)$. In this section, we will analytically show how the quality of the proposal distribution $\pi_{\text{prop}} = \pi_{\text{gen}}$ affects the quality of that estimator.

First, in Lemma A.2 we establish the variance of our estimator:

Lemma A.2. *For any LLM π_{gen} such that π_{ref} is absolutely continuous with respect to π_{gen} , the variance of the policy $\pi_{r,M}(\cdot|x)$ as an estimator for $\mathbb{E}_{y \sim \pi_r^*(\cdot|x)}[\pi_r^*(y|x)]$ is*

$$\text{Var}(\hat{\mu}_{\text{IS}}) = \frac{1}{M} \left(\sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{gen}}(y|x)} - \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2 \right). \quad (8)$$

Proof. Let y_1, y_2, \dots, y_M be the M generated candidate replies drawn from the proposal distribution $\pi_{\text{prop}}(\cdot|x)$. Conditioned on those candidate replies, the importance sampling estimator for the expected value of the function $\pi_r^*(y|x)$ under the target distribution $\pi_r^*(y|x)$ is

$$\hat{\mu}_{\text{IS}} = \frac{1}{M} \sum_{j=1}^M \pi_r^*(y_j|x) \cdot w(x_j) = \frac{1}{M} \sum_{j=1}^M \pi_r^*(y_j|x) \cdot \frac{\pi_r^*(y_j|x)}{\pi_{\text{prop}}(y_j|x)} = \frac{1}{M} \sum_{j=1}^M \frac{\pi_r^*(y_j|x)^2}{\pi_{\text{prop}}(y_j|x)}$$

Taking expectation over the generated candidate replies y_1, \dots, y_M :

$$\begin{aligned} \mathbb{E}[\hat{\mu}_{\text{IS}}] &= \mathbb{E}_{y_1, \dots, y_M \sim \pi_{\text{prop}}(\cdot|x)} \left[\frac{1}{M} \sum_{j=1}^M \frac{\pi_r^*(y_j|x)^2}{\pi_{\text{prop}}(y_j|x)} \right] \\ &= \frac{1}{M} \sum_{j=1}^m \mathbb{E}_{y_j \sim \pi_{\text{prop}}(\cdot|x)} \left[\frac{\pi_r^*(y_j|x)^2}{\pi_{\text{prop}}(y_j|x)} \right] \\ &= \mathbb{E}_{y \sim \pi_{\text{prop}}(\cdot|x)} \left[\frac{\pi_r^*(y|x)^2}{\pi_{\text{prop}}(y|x)} \right] \\ &= \sum_{y \in Y} \frac{\pi_r^*(y|x)^2}{\pi_{\text{prop}}(y|x)} \pi_{\text{prop}}(y|x) \\ &= \sum_{y \in Y} \pi_r^*(y|x)^2 \end{aligned} \quad (9)$$

Thus, our estimator is unbiased, as expected. The variance of the estimator $\hat{\mu}_{\text{IS}}$ is given by:

$$\text{Var}(\hat{\mu}_{\text{IS}}) = \mathbb{E}[\hat{\mu}_{\text{IS}}^2] - \mathbb{E}[\hat{\mu}_{\text{IS}}]^2 \quad (10)$$

For the term $\hat{\mu}_{\text{IS}}^2$ we have:

$$\hat{\mu}_{\text{IS}}^2 = \left(\frac{1}{M} \sum_{j=1}^M \frac{\pi_r^*(y_j|x)^2}{\pi_{\text{prop}}(y_j|x)} \right)^2 = \frac{1}{M^2} \sum_{j=1}^M \sum_{j'=1}^M \frac{\pi_r^*(y_j|x)^2}{\pi_{\text{prop}}(y_j|x)} \cdot \frac{\pi_r^*(y_{j'}|x)^2}{\pi_{\text{prop}}(y_{j'}|x)}$$

Taking the expectation $\mathbb{E}_{\pi_{\text{prop}}(\cdot|x)}[\cdot]$:

$$\begin{aligned}
\mathbb{E}[\hat{\mu}_{\text{IS}}^2] &= \frac{1}{M^2} \sum_{j=1}^M \sum_{j'=1}^M \mathbb{E}_{\pi_{\text{prop}}(\cdot|x)} \left[\frac{\pi_r^*(y_j|x)^2}{\pi_{\text{prop}}(y_j|x)} \cdot \frac{\pi_r^*(y_{j'}|x)^2}{\pi_{\text{prop}}(y_{j'}|x)} \right] \\
&= \frac{1}{M^2} \left(\sum_{j=1}^M \mathbb{E}_{\pi_{\text{prop}}(\cdot|x)} \left[\frac{\pi_r^*(y_j|x)^4}{\pi_{\text{prop}}(y_j|x)^2} \right] + \sum_{j \neq j'} \mathbb{E}_{\pi_{\text{prop}}(\cdot|x)} \left[\frac{\pi_r^*(y_j|x)^2}{\pi_{\text{prop}}(y_j|x)} \right] \right) \\
&= \frac{1}{M^2} \left(M \sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)} + M(M-1) \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2 \right) \\
&= \frac{1}{M} \sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)} + \frac{M-1}{M} \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2
\end{aligned} \tag{11}$$

The final expression for the variance of the importance sampling estimator with M samples can be computed by substituting Equations (9) and (11) in Equation (10):

$$\begin{aligned}
\text{Var}(\hat{\mu}_{\text{IS}}) &= \frac{1}{M} \sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)} + \frac{M-1}{M} \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2 - \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2 \\
&= \frac{1}{M} \left(\sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)} - \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2 \right)
\end{aligned} \tag{12}$$

□

Lemma A.2 reveals two important details about our estimator. First, the variance of the estimator is inversely proportional to the number of candidate replies generated M . Thus, as M increases the variance decreases and our estimator becomes more stable in simulating the target distribution $\pi_r^*(\cdot|x)$. Second, the closer that the proposal distribution $\pi_{\text{prop}}(\cdot|x)$ is to the target distribution, the smaller the term $\sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)}$ will be, which reduces the variance of our estimator.

Lemma A.3 establishes how a smaller variance leads to faster convergence:

Lemma A.3. *Let $\mu_{\text{true}} = \mathbb{E}_{y \sim \pi_r^*(\cdot|x)} [\pi_r^*(y|x)]$, and $\hat{\mu}_{\text{IS}}$ be the importance sampling estimator as in Lemma A.2 using $M = O\left(\frac{1}{\delta\varepsilon^2} \sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)}\right)$ candidate replies. Then, for any LLM π_{gen} such that π_{ref} is absolutely continuous with respect to π_{gen} , we have that $|\hat{\mu}_{\text{IS}} - \mu_{\text{true}}| < \varepsilon$ with probability at least $1 - \delta$.*

Proof. Let $\hat{\mu}_{\text{IS}}$ be the importance-weighted estimator whose variance is given by Lemma A.2 as

$$\text{Var}(\hat{\mu}_{\text{IS}}) = \frac{1}{M} \left(\sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)} - \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2 \right).$$

We have already shown in Lemma A.2 that the estimator is unbiased. By Chebyshev's inequality, for any $\varepsilon > 0$,

$$\Pr[|\hat{\mu}_{\text{IS}} - \mu_{\text{true}}| \geq \varepsilon] \leq \frac{\text{Var}(\hat{\mu}_{\text{IS}})}{\varepsilon^2} = \frac{1}{M \cdot \varepsilon^2} \left(\sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)} - \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2 \right)$$

where $\mu_{\text{true}} = \mathbb{E}_{y \sim \pi_r^*(\cdot|x)} [\pi_r^*(y|x)] = \sum_{y \in Y} \pi_r^*(y|x)^2$.

Setting $\Pr[|\hat{\mu}_{\text{IS}} - \mu_{\text{true}}| \geq \varepsilon] < \delta$ and solving for M yields:

$$M = \frac{1}{\delta\varepsilon^2} \left(\sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)} - \left(\sum_{y \in Y} \pi_r^*(y|x)^2 \right)^2 \right) = O\left(\frac{1}{\delta\varepsilon^2} \sum_{y \in Y} \frac{\pi_r^*(y|x)^4}{\pi_{\text{prop}}(y|x)}\right).$$

□

Based on Lemma A.3, our allocation rule converges to the quantity it is trying to approximate at a rate of \sqrt{M} .

A.3 BOUNDING THE TOTAL VARIATION DISTANCE

In Appendix A.1 we proved that MOSAIC converges in the limit to the theoretically optimal distribution for the platform's objective. Then, in Appendix A.2 we showed how MOSAIC's allocation rule relates to importance sampling, and what that implies for the similarity of MOISAC's output to that of the optimal distribution π_r^* . Now we will also show that MOSAIC's probability estimates for any possible sequence converge to that of π_r^* at a rate of \sqrt{M} .

Theorem A.4. *Let \mathcal{Y} be the domain consisting of all sequences in a finite token vocabulary T of length up to L . Hence $|\mathcal{Y}| < \infty$. Assume $0 \leq w(y) \leq C$ for all $y \in \mathcal{Y}$, where*

$$w(y) = \frac{\pi_r^*(y \mid x)}{\pi_{\text{prop}}(y \mid x)}.$$

Consider M i.i.d. draws $y_1, \dots, y_M \sim \pi_{\text{prop}}(\cdot \mid x)$ and define the self-normalized importance-sampled distribution induced by Algorithm 1:

$$\hat{\pi}_{r,M}(y \mid x) = \frac{\sum_{j=1}^M w(y_j) \mathbf{1}\{y_j = y\}}{\sum_{j=1}^M w(y_j)}.$$

Then, for any $\varepsilon > 0$ and $\delta > 0$, if $M = O\left(\frac{C^2 \cdot |\mathcal{Y}|^2}{\varepsilon^2} \ln\left(\frac{|\mathcal{Y}|}{\delta}\right)\right)$, with probability at least $1 - \delta$ (over the random draws), the Total Variation distance between $\hat{\pi}_M(\cdot \mid x)$ and $\pi_r^(\cdot \mid x)$ satisfies $d_{\text{TV}}(\hat{\pi}_M, \pi_r^*) \leq \varepsilon$.*

Proof. First, we will show that for any $M \in \mathbb{N}$, MOSAIC is an unbiased estimator of π_r^* .

For each fixed $y \in \mathcal{Y}$, define $X_j^{(y)} := w(y_j) \mathbf{1}\{y_j = y\}$ to be a random variable that equals $w(y_j)$ if the j -th draw from the proposal distribution is y . Intuitively, this is a weighted indicator for whether the j -th draw is equal to y .

For any $y \in \mathcal{Y}$, the expected value of this indicator, with the expectation taken over the draw of $y_j \sim \pi_{\text{prop}}$

$$\mathbb{E}[X_j^{(y)}] = \sum_{z \in \mathcal{Y}} w(z) \mathbf{1}\{z = y\} \pi_{\text{prop}}(z) = w(y) \pi_{\text{prop}}(y) = \pi_r(y).$$

Thus, $\hat{Q}_M(\{y\}) = \frac{1}{M} \sum_{j=1}^M X_j^{(y)}$ is an unbiased estimator of $\pi_r^*(y \mid x)$.

Similarly, $\hat{Q}_M(\mathcal{Y}) = \frac{1}{M} \sum_{j=1}^M w(y_j)$ is an unbiased estimator of 1 (since $\mathbb{E}[w(y_j)] = 1$) (assuming that π_r^* is the optimal policy, i.e., a normalized LLM so that the probabilities sum up to one).

The self-normalized distribution can be written as

$$\hat{\pi}_{r,M}(y \mid x) = \frac{\hat{Q}_M(\{y\})}{\hat{Q}_M(\mathcal{Y})}.$$

Hoeffding's inequality for bounded random variables tells us that for any $\varepsilon' > 0$:

$$\Pr\left[|\hat{Q}_M(\{y\}) - \pi_r^*(y \mid x)| \geq \varepsilon'\right] \leq 2 \exp\left(-\frac{2M\varepsilon'^2}{C^2}\right).$$

For each $y \in \mathcal{Y}$, we have the same deviation event, so by a union bound over all $y \in \mathcal{Y}$:

$$\Pr\left[\exists y \in \mathcal{Y} : |\hat{Q}_M(\{y\}) - \pi_r^*(y \mid x)| \geq \varepsilon'\right] \leq \sum_{y \in \mathcal{Y}} 2 \exp\left(-\frac{2M\varepsilon'^2}{C^2}\right) = 2|\mathcal{Y}| \exp\left(-\frac{2M\varepsilon'^2}{C^2}\right) \leq \delta_1.$$

Thus, using $M = O\left(\frac{C^2}{\varepsilon'^2} \ln\left(\frac{|\mathcal{Y}|}{\delta_1}\right)\right)$ candidate replies, with probability at least $1 - \delta_1$ over the draw of those replies, we have *simultaneously* for all $y \in \mathcal{Y}$, $|\hat{Q}_M(\{y\}) - \pi_r^*(y|x)| \leq \varepsilon'$.

By a similar argument, applying Hoeffding's inequality to $\hat{Q}_M(\mathcal{Y})$, which is an unbiased estimator of 1:

$$\Pr\left[|\hat{Q}_M(\mathcal{Y}) - 1| \geq \varepsilon'\right] \leq 2 \exp\left(-\frac{2M\varepsilon'^2}{C^2}\right).$$

Hence, with high probability $1 - \delta_2$, $\hat{Q}_M(\mathcal{Y})$ is also within $\pm\varepsilon$ of 1. A union bound over both events (all $y \in \mathcal{Y}$ plus $\hat{Q}_M(\mathcal{Y})$) yields a final event E'_ε with probability $1 - (\delta_1 + \delta_2)$ on which:

$$\hat{Q}_M(\{y\}) \in [\pi_r^*(y) - \varepsilon', \pi_r^*(y) + \varepsilon'], \quad \hat{Q}_M(\mathcal{Y}) \in [1 - \varepsilon', 1 + \varepsilon'], \quad \forall y \in \mathcal{Y}.$$

On this good event E'_ε , for each $y \in \mathcal{Y}$:

$$\hat{\pi}_M(y|x) = \frac{\hat{Q}_M(\{y\})}{\hat{Q}_M(\mathcal{Y})} \in \left[\frac{\pi_r^*(y|x) - \varepsilon'}{1 + \varepsilon}, \frac{\pi_r^*(y|x) + \varepsilon'}{1 - \varepsilon'}\right]$$

Recall the total variation (TV) distance between two discrete distributions P and Q on \mathcal{Y} is

$$d_{\text{TV}}(P, Q) = \frac{1}{2} \sum_{y \in \mathcal{Y}} |P(y) - Q(y)|.$$

We aim to show that on the event E'_ε (which holds with probability at least $1 - (\delta_1 + \delta_2)$), the quantity

$$\sum_{y \in \mathcal{Y}} |\hat{\pi}_M(y|x) - \pi_r^*(y|x)|$$

is at most $O(\varepsilon')$.

On E'_ε , we have, for each $y \in \mathcal{Y}$,

$$\hat{\pi}_M(y|x) = \frac{\hat{Q}_M(\{y\})}{\hat{Q}_M(\mathcal{Y})} \in \left[\frac{\pi_r^*(y|x) - \varepsilon'}{1 + \varepsilon'}, \frac{\pi_r^*(y|x) + \varepsilon'}{1 - \varepsilon'}\right].$$

Thus:

$$\hat{\pi}_M(y|x) - \pi_r^*(y|x) \in \left[-\frac{\varepsilon' \pi_r^*(y|x)}{1 + \varepsilon'} - \frac{\varepsilon'}{1 + \varepsilon'}, \frac{\varepsilon' \pi_r^*(y|x)}{1 - \varepsilon'} + \frac{\varepsilon'}{1 - \varepsilon'}\right] \quad (13)$$

$$\in \left[-\frac{2\varepsilon}{1 + \varepsilon}, \frac{2\varepsilon}{1 - \varepsilon}\right] \quad (14)$$

Thus, conditioned on the good event E'_ε , we have $|\hat{\pi}_M(y) - \pi_r^*(y)| = O(\varepsilon')$ by a Taylor expansion for small ε' .

Therefore, on the event E'_ε ,

$$d_{\text{TV}}(\hat{\pi}_M, \pi_r^*) = \frac{1}{2} \sum_{y \in \mathcal{Y}} |\hat{\pi}_M(y) - \pi_r^*(y)| \leq \frac{1}{2} \sum_{y \in \mathcal{Y}} O(\varepsilon') = O(|\mathcal{Y}| \cdot \varepsilon').$$

Setting $\varepsilon' = \frac{\varepsilon}{|\mathcal{Y}|}$ to achieve the target error ε concludes the proof. \square

B DETAILS FROM SECTION 5

In this section, we present all omitted details from Section 5.

B.1 OMITTED PROOFS FROM SECTION 5.1

Theorem 5.1 Proof. Let $\hat{\mathbf{r}}_i = (\hat{r}_i(x, y_1), \dots, \hat{r}_i(x, y_M))$ be the reward reports of advertiser i for the M generated candidate sequences. Then, for MOSAIC's allocation rule, holding the candidate sequences and the reports of all other advertisers fixed, the probability of returning each of the M candidate sequences as a function of i 's reports is:

$$\pi(\mathbf{r}_i; \vec{\beta}_{-i}) = \text{softmax} \left(\frac{\mathbf{r}_i}{\tau} + \boldsymbol{\beta}_{-i} \right), \quad (15)$$

where $\boldsymbol{\beta}_{-i,j} = \frac{\sum_{k \in N \setminus \{i\}} \hat{r}_k(x, y_j)}{\tau} + \log \frac{\pi_{\text{ref}}(y_j|x)}{\pi_{\text{con}}(y_j|x;c)}$. Importantly $\boldsymbol{\beta}_{-i}$ is an M -dimensional vector that does not depend on advertiser i 's reports.

We would like to equip $\pi(\cdot; \boldsymbol{\beta}_{-i})$ with a payment rule $p(\cdot; \boldsymbol{\beta}_{-i})$ so that the resulting mechanism $(\pi(\cdot; \boldsymbol{\beta}_{-i}), p(\cdot; \boldsymbol{\beta}_{-i}))$ will be strategyproof. This requires that $\pi(\cdot; \boldsymbol{\beta}_{-i})$ have a property known as *cyclic monotonicity*. Equivalently, $\pi(\cdot; \boldsymbol{\beta}_{-i})$ must be the (sub)gradient of advertiser i 's utility for bidding truthfully in the mechanism $U(\mathbf{r}_i; \boldsymbol{\beta}_{-i})$, and that utility function must be convex (Frongillo & Kash, 2021; Rochet, 1987; Myerson, 1981).

It is easy to verify that for the function class:

$$U_C(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) = \tau \log \left(\sum_{j=1}^M \exp \left(\frac{r_i(x, y_j)}{\tau} + \boldsymbol{\beta}_{-i,j} \right) \right) + C, \quad C \in \mathbb{R} \quad (16)$$

the allocation rule $\pi(\mathbf{r}_i; \boldsymbol{\beta}_{-i})$ is a gradient of $U_C(\mathbf{r}_i; \boldsymbol{\beta}_{-i})$. Additionally, $U_C(\mathbf{r}_i; \boldsymbol{\beta}_{-i})$ is convex in \mathbf{r}_i : The exponential function e^x is (strictly) convex, because its second derivative is positive. The transformation $\frac{r_i(x, y_j)}{\tau} + \boldsymbol{\beta}_{-i,j}$ is an affine transformation of $r_i(x, y_j)$, and affine transformations preserve convexity. Finally, it is well-known that the LogSumExp function is convex.

Thus, for any $\boldsymbol{\beta}_{-i}$ and for any set of generated candidate sequences, reporting truthfully maximizes advertiser i 's expected utility, with the expectation taken over the draw of the final sequence from the set of candidate sequences. Adopting the quasi-linear utility model, advertiser i 's payment is:

$$\begin{aligned} U_C(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) &= \pi(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) \cdot \mathbf{r}_i - p(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) \\ p(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) &= \pi(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) \cdot \mathbf{r}_i - U_C(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) \\ p(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) &= \pi(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) \cdot \mathbf{r}_i - \tau \log \left(\sum_{j=1}^M \exp \left(\frac{r_i(x, y_j)}{\tau} + \boldsymbol{\beta}_{-i,j} \right) \right) - C, \quad C \in \mathbb{R} \end{aligned} \quad (17)$$

□

B.2 DIFFERENCES FROM STANDARD AUCTION SETTINGS

Standard auction environments typically rely on a set of assumptions that simplify mechanism design; however, these assumptions do not apply to auctions for LLM-generated content. In this section, we detail these assumptions and discuss why they are inapplicable in our context.

First, in a standard auction setting, it is common to assume that the agents' valuation functions satisfy free disposal, i.e., $v_i(S) \geq v_i(S') \forall S \supseteq S', S, S' \subseteq \mathcal{I}$. The interpretation of free disposal is that an agent can discard any items she is allocated that she is not interested in. Free disposal combined with the fact that an agent has zero value for the empty bundle mean that her value for any outcome is weakly positive. Second, in most auction environments, the allocation rule is different for different agents: each agent will get allocated her own bundle of items, and we can assume that she is indifferent to the allocation of items to the other agents.

As detailed in [Rafailov et al. \(2023\)](#), assuming that an agent's LLM π_i was trained to maximize her reward function (and regularized with respect to its KL divergence from some reference LLM, which we assume to be the same as the auctioneer's reference LLM), there is a one-to-many mapping between an advertiser's optimal LLM, and her implicit reward function. That mapping is:

$$r_i(x, y) = \tau_i \log \frac{\pi_i(y|x)}{\pi_{\text{ref}}(y|x)} + \log Z_i(x) \quad (18)$$

where $Z_i(x)$ is a prompt-dependent constant, and τ_i is the regularization hyperparameter of advertiser i , similar to the one in Equation (1). All functions in the class defined in Equation (18) are equivalent, in the sense that they induce exactly the same LLM ([Rafailov et al., 2023](#)). This has two implications: First, unlike standard auction environments, an agent's reward can go negative – there is nothing equivalent to the free disposal property. Setting $Z_i(x)$ to zero (which is equivalent to normalizing the induced probabilities by the LLM ([Rafailov et al., 2023](#))), the agent's reward is negative for any sequence for which her LLM assigns a lower probability than π_{ref} .

Second, especially in the online advertising application, an agent's expected utility for not participating in the auction is negative: if advertiser i does not participate in the auction, her payment is zero, but her expected value for the outcome is

$$\pi(\mathbf{0}; \boldsymbol{\beta}_{-i}) \cdot \mathbf{r}_i = \pi(\boldsymbol{\beta}_{-i}) \cdot \mathbf{r}_i \quad (19)$$

The other advertisers have very low rewards for the sequences that mention advertiser i : assuming their LLMs have been properly trained, they will evaluate all sequences that explicitly mention a different, possibly competing brand, as unlikely. Thus, based on Equation (18) the corresponding advertisers have very low rewards for those sequences and conversely, advertiser i has low rewards for the sequences that the other advertisers have high rewards for. But based on Equation (2), if advertiser i does not participate in the auction, $\pi(\mathbf{0}; \boldsymbol{\beta}_{-i})$ will assign high probabilities to sequences for which i has low rewards for. Thus, Equation (19) implies that, unlike standard auction environments, the advertiser's expected reward and utility for not participating in the mechanism is negative.

B.3 OUR MECHANISM IS “ALMOST INDIVIDUALLY RATIONAL”

First, we explain why the standard notion of individual rationality (i.e., weakly positive utility from participation in the mechanism) encountered in most auction settings is impossible to achieve in this domain while converging to the optimal distribution and maintaining incentive compatibility. Then, we explain how, with our payment offset, our mechanism is “almost IR.” In Lemma B.1 we prove that the ex-interim utility of an advertiser who has zero reward for all candidate sequences and bids truthfully is deterministically zero, i.e., advertisers that do not contribute to the social welfare (but also do not detract from it) have zero utility. Similarly, in Lemma B.2 we prove that if an agent’s reward for all candidate sequences is (weakly) positive, then her ex-interim utility is (weakly) positive.

Why is individual rationality (IR) impossible? *Individual rationality* (IR) stipulates that an agent gains more utility by participating and bidding truthfully in a mechanism than by not participating at all. Typically, if an agent’s utility for non-participation is zero, participating should yield weakly positive utility. However, this simplification does not apply in our setting.

As discussed in Section 5.2, advertiser i ’s reward for any sequence y can be arbitrarily negative (Equation (18)). The same is true for the utility from truthful participation, as outlined in Equation (16). To ensure a positive utility for every advertiser in our mechanism, an offset would need to be infinitely large or dependent on advertiser i ’s reports. But then the mechanism’s allocation rule would no longer be the gradient of advertiser i ’s utility with respect to her reports, which would destroy strategyproofness (Frongillo & Kash, 2021; Rochet, 1987; Myerson, 1981).

It is important to note that this challenge is inherent not just to our mechanism but to any mechanism in this setting that operates with a fixed set of sequences, aims to approximate the optimal distribution, and maintains strategyproofness. Under these conditions, the only allocation rule that approximates the theoretically optimal distribution (Equation (2)) is that of our mechanism. However, this uniquely determines the advertisers’ utilities, up to a constant factor, as described in Equation (16) (Frongillo & Kash, 2021; Rochet, 1987; Myerson, 1981).

Lemma B.1. *For the payment offset $C = -\tau \log \left(\sum_{j=1}^M \exp(\beta_{-i,j}) \right)$ if advertiser i ’s reward for all candidate sequences is zero, then her ex-interim utility is deterministically zero, for all $\beta_{-i} \in \hat{R}_{-i}$.*

Lemma B.1 Proof. First, note that for all $\beta_{-i} \in \hat{R}_{-i}$, advertiser i ’s expected reward for the outcome is zero, as $\pi(r_i; \beta_{-i}) \cdot r_i = \pi(r_i; \beta_{-i}) \cdot \mathbf{0} = 0$. Additionally, advertiser i ’s reward for the realized outcome will deterministically be zero, as her reward for all generated candidate sequences is zero. Finally, note that by setting $r_i = \mathbf{0}$ in Equation (17) with the offset C set as in Section 5.3, we have that the advertiser i ’s payment is also deterministically zero. Thus, an advertiser with zero reward for all generated candidate sequences who reports her rewards truthfully has deterministically zero reward for the final outcome and zero payments, and her utility is also deterministically zero.

□

Lemma B.2. *For the payment offset $C = -\tau \log \left(\sum_{j=1}^M \exp(\beta_{-i,j}) \right)$ if advertiser i ’s reward for all candidate sequences is positive, then her ex-interim utility is positive, for all reports $\beta_{-i} \in \hat{R}_{-i}$.*

Proof. Lemma B.1 establishes that when advertiser i ’s reward for all candidate sequences is zero, her utility for truthfully bidding in the mechanism, denoted as $U(\mathbf{0}; \beta_{-i})$, is zero for all possible reports of the other advertisers $\beta_{-i} \in \hat{R}_{-i}$.

Furthermore, Theorem 5.1 shows that the mechanism’s allocation rule corresponds to the gradient of advertiser i ’s utility when bidding truthfully. Because the allocation rule is non-negative, the gradient of advertiser i ’s utility for bidding truthfully is also non-negative.

Thus, if advertiser i ’s rewards for all candidate sequences are weakly positive, and considering the non-negative gradient of her utility, her ex-interim utility under truthful bidding must be positive, irrespective of the other advertisers’ reports β_{-i} .

□

Corollary B.3. For the payment offset $C = -\tau \log \left(\sum_{j=1}^M \exp(\beta_{-i,j}) \right)$ if the distribution π_{con} only generates candidate sequences for which advertiser i 's reward is positive, then the ex-ante expected utility of the advertiser is positive.

Corollary B.3 Proof. This follows immediately from the fact the the fact that the ex-ante utility of the advertiser is the expectation of her ex-interim utility with respect to her reward for the generated sequences, and the fact that the second quantity is positive whenever the reward of the advertiser for all candidate sequences is positive from Lemma B.2.

□

B.4 “WHAT YOU GIVE IS WHAT YOU GET”

As we explained in Section 5.3.2, our allocation rule, which is the only one over a finite set of replies that converges to the optimal LLM, is also the (sub)gradient of the utility to ensure truthfulness (Rochet, 1987). Because the allocation rule is the same for all advertisers, their utilities must also be the same, up to advertiser-specific offsets, as indicated by Equation (20):

$$U_C(\mathbf{r}_i; \boldsymbol{\beta}_{-i}) = \tau \log \left(\sum_{j=1}^M \exp \left(\frac{1}{\tau} \sum_{k \in N} r_k(x, y_j) \right) + \log \frac{\pi_{\text{ref}}(y_j|x)}{\pi_{\text{con}}(y_j|x; c)} \right) + C, \quad C \in \mathbb{R} \quad (20)$$

However, not all advertisers contribute equally to the social welfare of the final outcome. Because of this, implementing the mechanism without a carefully-designed offset would lead to free-riding: as long as an agent's utility in Equation (20) is positive, she would be incentivized to participate, even if the user query was completely unrelated to her business, because the mechanism would ensure that she received, on expectation, the same (positive) expected utility from doing so as any other participating advertiser.¹⁰

Incentivizing unrelated advertisers to participate would have adverse effects. First, the better-performing context-aware mechanism would create candidate sequences with worse rewards for *all* advertisers, because its context would be “diluted” from advertisers unrelated to the user query. In our running example for the query “How to learn music online?”, imagine adding “Try to mention ‘EasySwitch’, a comprehensive VPN service” to the context of the context-aware LLM $\pi_{\text{con}}(\cdot|x; c)$.

Additionally, for both versions of the mechanism, following the discussion in Section 5.2, the advertisers for whom the user query is unrelated are more likely to have negative rewards for the generated sequences as their LLMs will deem the candidate sequences more unlikely than the reference LLM. Thus, based on Equation (5), their participation in the mechanism will lead to a reduction of the total sum of rewards of the generated sequences for the advertisers, which will indirectly reduce the expected utility of all advertisers, making the mechanism less attractive for the user-query-relevant advertisers.

To summarize, all advertisers receiving the same utility would incentivize advertisers for whom the user query is unrelated to participate in the auction. This would in turn reduce everyone's expected utility, potentially reducing the incentive for the user-query-relevant advertisers to participate, and lead to sequences with worse expected rewards for the advertisers and usefulness for the user.¹¹ Thus, in the application of auctions for aggregating advertisers' preferences over LLM-generated outputs, advertisers with higher contribution to social welfare also receiving proportionally higher utility by the mechanism is important for the long-term success of the mechanism in practice.

¹⁰We can assume that advertisers can estimate their expected utility from participation using historical data from past auctions, analogously to how they can estimate their utility for participating in sponsored search auctions.

¹¹If we interpret the KL divergence between the distribution induced by the reference LLM and the LLM that generated the candidate sequences as a measure of their expected usefulness for the user.

C DETAILS FROM SECTION 6

C.1 DETAILED EXPERIMENT SETUP

We create a set of synthetic instances to test our mechanism. Each instance consists of a user query, e.g. “How do I bake cookies?” and a list of advertisers. Each advertiser is defined by an “advertiser name”, e.g. “KitchenFix” and an advertiser description, e.g., “producing kitchen appliances.”¹² The reference LLM π_{ref} responsible for generating replies that are useful for the user is Llama-2-7b-chat-hf (Touvron et al., 2023). The advertisers’ LLMs are created using the same reference LLM, and adding the instruction: “Answer the question advertising ⟨advertiser⟩, ⟨advertiser description⟩.” The context aware LLM is created using the same reference LLM, and adding the instruction: “Answer the query. Try to mention ⟨advertiser 1⟩, who ⟨advertiser description 1⟩ and ⟨advertiser 2⟩, who ⟨advertiser description 2⟩.”

Following (Rafailov et al., 2023) the reward function of advertiser i is set to $r_i(x, y) = \log \frac{\pi_i(y|x)}{\pi_{\text{ref}}(y|x)}$, where π_i is advertiser i ’s LLM, i.e., we set $\tau_i = 1$, $Z_i(x) = 1$ for all advertisers and for all user prompts in Equation (18).¹³ For the auctioneer’s objective as defined in Equation (1) we set $\tau = 1$, balancing between the advertisers’ expected rewards for the generated sequences and the sequences’ divergence from the reference LLM responsible for generating useful replies for the user. Thus, the optimal policy according to Equation (2) becomes:

$$\pi_r^*(y|x) = \frac{1}{Z(x)} \pi_{\text{ref}}(y|x) \exp \left(\sum_{i \in N} r_i(x, y) \right) \quad (21)$$

Following Li et al. (2024); Rozière et al. (2024) we sample from the LLM generating the sequences (either π_{ref} or π_{con}) with temperature 0.8 and top-p 0.95. We use 50 user queries, each with two interested advertisers. To increase the statistical significance of our results, we test each user query on 25 different random seeds, effectively resulting in 1250 individual instances. The seed affects both the generation of the candidate sequences, and the randomized sampling of the final sequence from the candidate sequences. We will make all of our code and set of synthetic instances publicly available.

All experiments were conducted on a compute cluster running Ubuntu 20.04.6 LTS with AMD EPYC processors with 48 cores and 1512GB RAM and Nvidia A100 GPUs and Python 3.12.1.

¹²Following Düting et al. (2024), we use fictitious brand names to avoid the model using any existing ad.

¹³Note that we cannot estimate the true value of $Z_i(x)$. However, it is a query-dependent constant. Given that our experiments focus on the reward difference of different sequences for the same query, this is a non-issue.

C.2 COMPARISON OF BASELINE AND CONTEXT-AWARE MECHANISM OUTPUTS

To provide some intuition of how the context-aware mechanism improves the relevance of generated responses for the advertisers, we present a side-by-side comparison of outputs from both mechanisms for a specific query in Table 2. Note that only the context-aware mechanism is able to incorporate the advertisers into the reply to the user’s query.

User query
What are effective ways to learn a musical instrument online?
Advertisers
- MusicMastery: offering online music lessons and virtual tutoring - InstaTune: selling musical instruments and learning aids
Context-Aware Mechanism Output
Effective ways to learn a musical instrument online include: 1. Online music lessons: Platforms like MusicMastery offer one-on-one virtual tutoring, group classes, and pre-recorded lessons. These lessons can be tailored to your skill level and learning style. 2. Interactive learning tools: InstaTune offers a range of interactive learning tools, such as apps, software, and games, that make learning fun and engaging. 3. Practice with technology: Utilize digital tools like metronomes, tuners, and recording software to practice and track your progress. 4. Collaborate with other musicians: Join online communities, forums, and social media groups to connect with other musicians and learn from their experiences. ...
Baseline Mechanism Output
- Practice consistently and regularly: Consistent practice is essential to learning any musical instrument, and it's even more important when learning online. Set aside a specific time each day or week to practice, and make sure to stick to it. - Find a good teacher or mentor: While it's possible to learn some things on your own, working with a good teacher or mentor can help you improve faster and avoid developing bad habits. Look for someone who is experienced and knowledgeable in the type of music you want to learn. - Use technology to your advantage: There are many online tools and resources available that can help you learn an instrument. These include video lessons, online tutorials, and apps that can help you practice and track your progress. ...

Table 2: Comparison of context-aware and baseline mechanism outputs using 10 candidate sequences. Advertisers highlighted in bold, outputs truncated to conserve space.

C.3 EXPERIMENTAL EVALUATION OF THE USEFULNESS TO THE USER

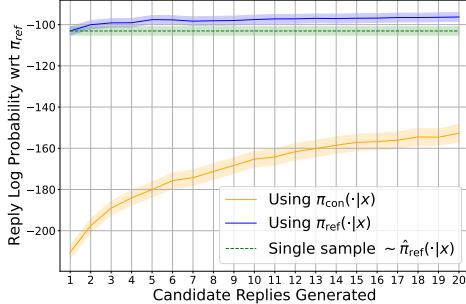


Figure 3: Reply log probability with respect to the reference LLM as a function of the number of replies generated using π_{ref} and π_{con} .

In this section, we experimentally evaluate the usefulness of the replies returned by MOSAIC from the user’s perspective.

In Figure 3, we plot the log probability of the replies returned by MOSAIC with respect to the optimal LLM for the user (i.e., the reference LLM) as a function of the number of candidate replies generated. We compare the context-aware version of MOSAIC to a baseline version. To benchmark our results, we also compare against sampling a *single reply* directly from the optimal distribution for this task, which is the reference LLM itself.

The results in Figure 3 show that for the context-aware version of our mechanism, the log probability of the replies relative to the reference LLM increases rapidly as the number of candidate replies grows. This demonstrates that by generating more replies, the context-aware version of MOSAIC is able to produce replies with substantial value for the user. By comparison, the usefulness to the user that the baseline version is able to achieve does not scale with more generated replies.

However, there remains a gap between the log probability of those replies and the benchmark set by the reference LLM. This discrepancy arises because the platform optimizes a different objective: the expected advertiser rewards combined with the KL divergence from the reference LLM. By increasing the weight τ that the platform places on the reference LLM, this gap between the context-aware version of MOSAIC and the user’s optimal benchmark will close.

C.4 COMPREHENSIVE EXPERIMENTAL EVALUATION OF THE OFFSET FROM SECTION 5.3

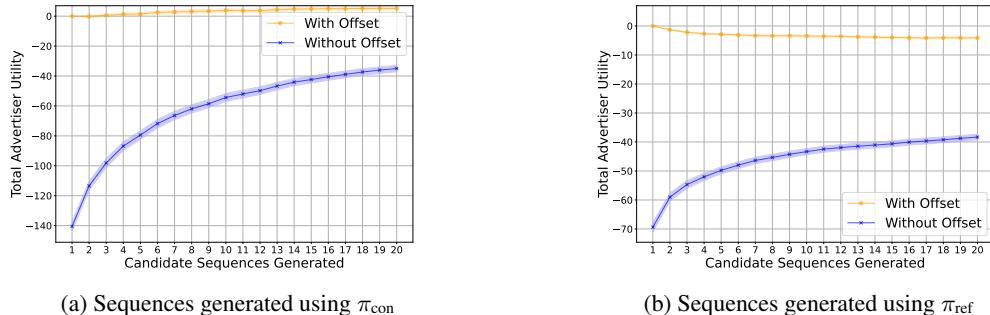


Figure 4: Comparison of total advertiser utility gain from participation with, and without the payment offset, as a function of the number of candidate sequences generated using π_{ref} and π_{con} . Averaged over 1250 runs including 95% CIs.

In this section, we provide a comprehensive experimental evaluation of the payment offset introduced in Section 5.3. Figure 4 explores the effectiveness of the payment offset introduced in Section 5.3

in ensuring that our mechanism is ex-ante IR, i.e., advertisers have positive expected utility gains from participating. To that end, we compare the advertiser utility gain from participation with and without the payment offset, as a function of the number of candidate sequences generated by the context-aware (Figure 4a) and baseline (Figure 4b) versions of our mechanism.

In Figure 4a we observe that for the context-aware version of our mechanism, adding the payment offset introduced of Section 5.3 to the payment rule is enough to ensure positive expected utility for the advertisers (conditioned on the fact that they are related to the user’s query), i.e., make the mechanism ex-ante IR. Notably, without the payment offset, advertisers have very negative expected utility from participating in the mechanism.

In Figure 4b we observe that for the baseline version of our mechanism, adding the payment offset introduced of Section 5.3 to the payment rule causes a very large increase in the advertisers’ expected utility, but it is still not enough to ensure positive expected utility for the advertisers. To conclude, Figure 4 demonstrates that the offset introduced in Section 5.3 causes a large increase in the advertisers’ expected utility, which in case of the context-aware version of the mechanism, is also enough to make the mechanism ex-ante IR.

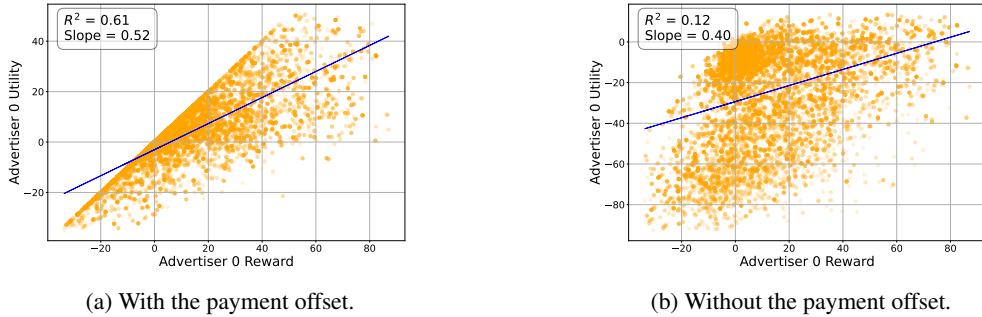


Figure 5: Comparative scatter plots of advertiser reward and utility gain from participation, with and without the payment offset of Section 5.3 for candidate sequences generated by the context-aware LLM π_{gem} . We additionally show a linear regressor fit to that data, its slope and its R^2 .

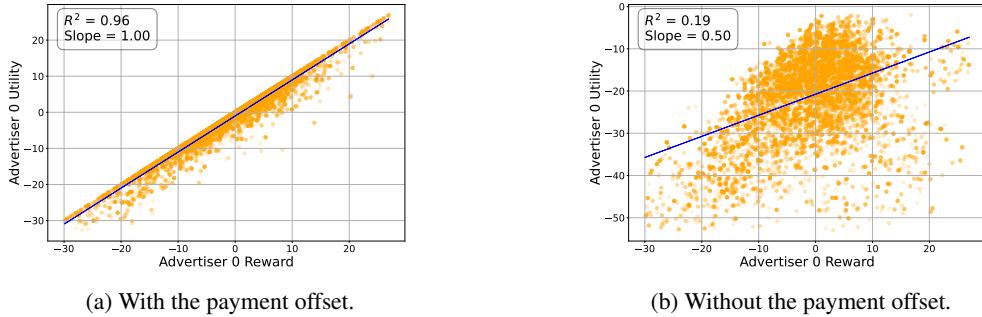


Figure 6: Comparative scatter plots of advertiser reward and utility gain from participation, with and without the payment offset of Section 5.3 for candidate sequences generated by the reference π_{ref} . We additionally show a linear regressor fit to that data, its slope and its R^2 .

Figures 5 and 6 explore the effectiveness of the payment offset introduced in Section 5.3 in aligning an advertiser’s utility with her contribution to the social welfare. In Figure 5 we compare the scatter plots of the advertiser reward and utility gain from participation in the mechanism, with and without the payment offset introduced in Section 5.3 for candidate sequences generated using the context-aware LLM π_{ref} . Additionally, for both subfigures, we show a linear regressor fitted to the data, as well as its slope and coefficient of determination. Comparing the two subfigures, it is immediately obvious that adding the offset to the payments makes the relationship between advertiser reward and utility gain far more linear. This is confirmed by the coefficient of determination of the linear regressors

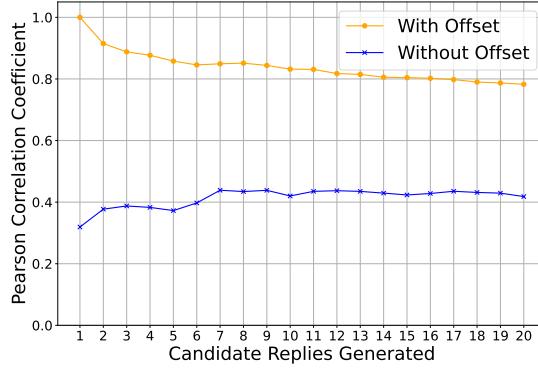


Figure 7: Pearson correlation between advertiser normalized utility and reward.

fit to each dataset. The coefficient of determination of the linear regressor is far larger when we use the offset. Without the payment offset, the coefficient of determination is almost 0, indicating that, without our payment offset, reward gain is not a predictive measure of an agent’s utility. Additionally, the slope of the linear regressor is also higher for the scatter plot with the payment offset.

In Figure 6 we make the same comparison, but for candidate sequences generated using the reference LLM π_{ref} . The results are now even more pronounced. In Figure 6a we observe the relationship between advertiser utility and reward gain with our payment offset is almost perfectly linear, as suggested by the linear regressor fitted to the data having a slope of 1.00 and an extremely high coefficient of determination of 0.96, indicating that it can almost perfectly fit the data. Without our payment offset however, in Figure 6b we can see that the relationship between the two metrics is again both less linear, and less positively correlated, as the slope of the linear regressor is 0.5 and its coefficient of determination is only 0.19.

To conclude, in all cases tested, the use of the advertiser-specific offset introduced in Section 5.3 increases an advertiser’s expected utility, makes the relationship between an advertiser’s contribution to social welfare and her utility both more linear and more positively correlated, and, for the context-aware version of our mechanism, can ensure positive expected utility for the advertisers.

C.5 EXPERIMENTS IN SETTINGS WITH MANY ADVERTISERS

In this section, we test MOSAIC’s performance in settings with $n = 5$ and $n = 10$ advertisers, compared to the results of Section 6 where the number of advertisers was set to 2. To maintain the same balance between advertisers and the usefulness to the user as in the experiments of the main paper, we set $\tau = n/2$.¹⁴ We use the same 50 user queries, but now each query has five and ten advertisers, and test each query on 20 different random seeds, resulting in 1000 instances.

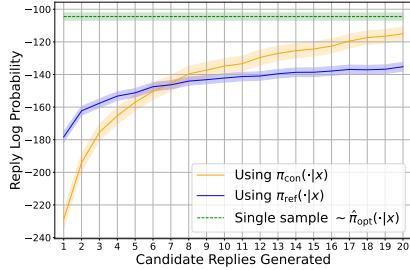
In Figures 8a to 8j, we present the log probability of the returned reply with respect to the optimal LLM for the platform’s objective π_r^* , and the reference LLM π_{ref} , the total advertiser value and reward gain from participating in MOSAIC and the platform’s revenue. All plots are with respect to the number of candidate replies (i.e., LLM queries) that MOSAIC used. We compare MOSAIC’s baseline and context-aware versions.

First, in Figures 8a and 8b we observe that the context-aware version of MOSAIC quickly converges to the platform’s objective. By comparison, MOSAIC’s naive version fails to do so. It is noteworthy that for very low numbers of candidate replies, the naive version of the mechanism performs better than the context-aware version. That is because in these experiments, more weight is placed towards the reference LLM, as we have set a higher τ value. However, the superior convergence rate of the context-aware version of the mechanism recovers that difference very quickly. At the same time, Figures 8c and 8d show that the context-aware version of the mechanism produces sequences that closely match the distribution of the reference LLM, thus maintaining the usefulness to the user.

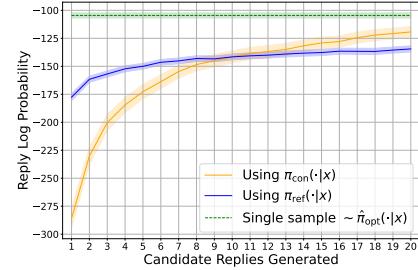
Finally, the context-aware version of MOSAIC is able to generate significant value (Figures 8e and 8f) and utility (Figures 8g and 8h) for the advertisers, while also recapturing a significant portion as revenue for the platform (Figures 8i and 8j).

Taken together, similar to our results in the main paper, we have shown that even with a large number of advertisers, MOSAIC quickly converges to the theoretically optimal distribution, generating significant value and utility for the advertisers and revenue for the platform, while also maintaining its usefulness to the user.

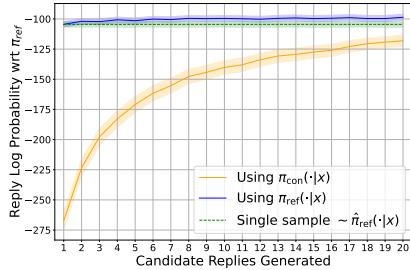
¹⁴Note that in the main experiments of Section 6 we had $n = 2$ and $\tau = 1$



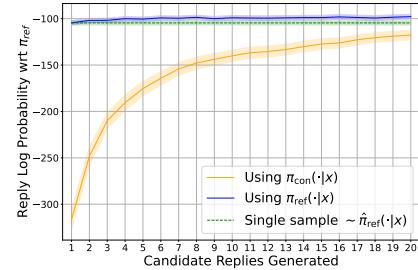
(a) Log probability w.r.t. π_r^* for 5 advertisers.



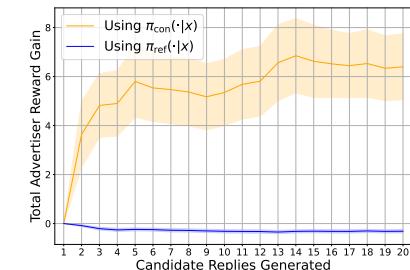
(b) Log probability w.r.t. π_r^* for 10 advertisers.



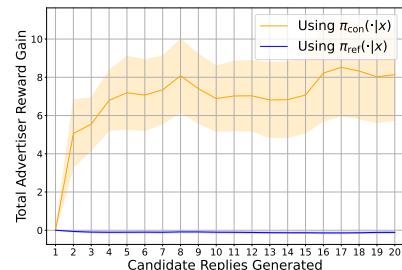
(c) Log probability wrt. π_{ref} for 5 advertisers.



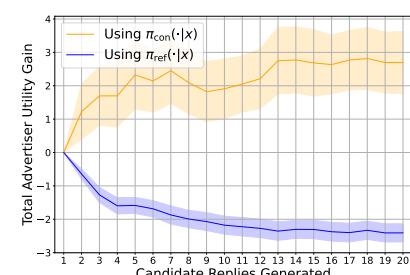
(d) Log probability wrt. π_{ref} for 10 advertisers.



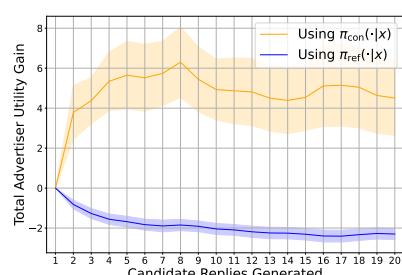
(e) Advertiser value gain for 5 advertisers.



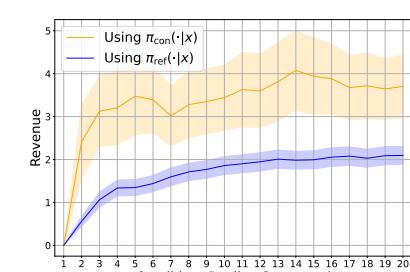
(f) Advertiser value gain for 10 advertisers.



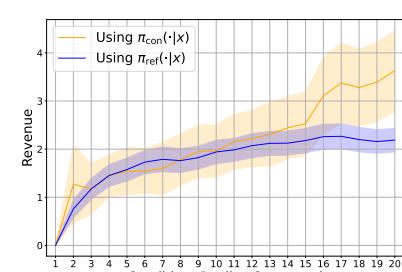
(g) Advertiser utility gain for 5 advertisers.



(h) Advertiser utility gain for 10 advertisers.



(i) Revenue for 5 advertisers.



(j) Revenue for 10 advertisers.

Figure 8: Comparison of results for 5 advertisers (left) and 10 advertisers (right). Each row corresponds to a specific metric: log probability with respect to the optimal LLM π_r^* , the reference LLM π_{ref} , advertiser value and utility gain, and revenue. Shown are averages over 1000 instances including 95% CIs.

C.6 COMPUTE EXPERIMENTS

In this section, we experimentally evaluate MOSAIC’s computational requirements. Using the setup described in Section 6.1, we measure the total wall time required by MOSAIC on a *single* A100 GPU, focusing on the number of candidate replies generated for evaluation. This total time includes:

1. Generating candidate replies using the context-aware LLM.
2. Evaluating these replies with the advertisers’ LLMs implemented as discussed in Section 6.1.
3. Calculating the Rochet payments, as detailed in Section 5. In practice the time to calculate payments is negligible (under 10 milliseconds), and need not increase user-perceived latency because they can be calculated after showing outputs to the user.

Figure 9 presents the total time taken by MOSAIC, which includes both the generation and evaluation of candidate replies. These components are further broken down in Figures 10 and 11, respectively.

As previously demonstrated in Section 6 and Appendix C.5, MOSAIC converges to the optimal distribution using only 20 candidate replies, regardless of the number of advertisers. In Figure 10, we observe that generating 10 and 20 candidate replies on a single A100 GPU takes 30.5 and 60.5 seconds, respectively. This represents a 2.5- and 5-fold increase over the 12.0 seconds required to generate a single reply from the same LLM, i.e., the latency perceived by the user when directly querying the LLM.

To reduce user-perceived latency, MOSAIC can parallelize candidate generation across multiple GPUs. For example, distributing the generation of 20 replies across two GPUs (each generating 10 replies) reduces the total time to 30.5 seconds, assuming efficient evaluation techniques as outlined in Section 4.2. This approach results in MOSAIC requiring approximately 5 times the compute time of generating a single reply to converge, but with a perceived latency to the user of only 2.5 times that of directly querying an LLM (of the same architecture and similar size). Although our experiments were run on a single GPU, note that MOSAIC can always be parallelized further so that the user-perceived latency is the same as directly querying a *single* LLM, as explained in Section 4.2.

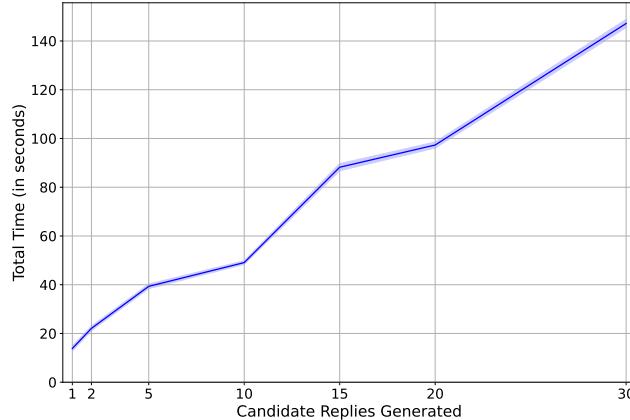


Figure 9: MOSAIC total wall time (seconds) as a function of the number of candidate replies generated using π_{con} . Shown are averages over 50 instances including 95 % CIs.

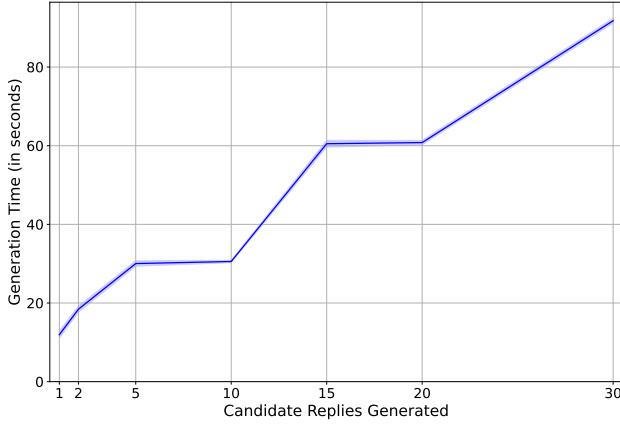


Figure 10: MOSAIC generation time as a function of the number of candidate replies generated using π_{con} . Shown are averages over 50 instances including 95 % CIs.

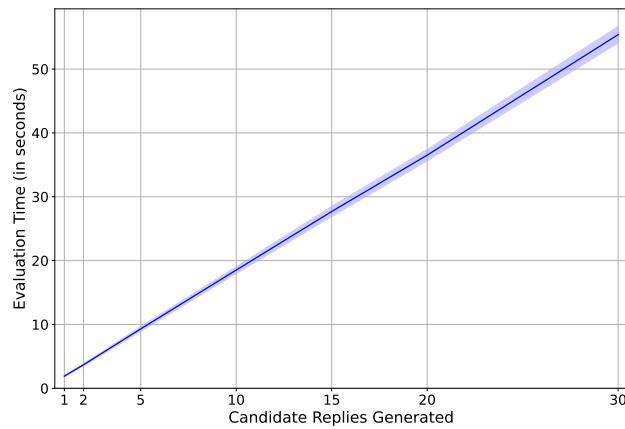


Figure 11: MOSAIC evaluation time as a function of the number of candidate replies. Shown are averages over 50 instances including 95 % CIs.

C.7 CONFLICTS BETWEEN ADVERTISERS

The output of MOSAIC is a single text response to the user. In trying to reflect the interests of multiple advertisers, there is a risk of a single response being incoherent, just mentioning as many advertisers as possible without being a good promotion for any of them. However, the platform's objective, which takes into account both advertiser utility and usefulness for the user, should account for this: if advertisers are happy to be cross-promoted, those outcomes should be more likely, but incoherent advertisements or those that simultaneously promote competitors should be less likely.

In Figure 12 we show the fraction of advertisers mentioned by the context-aware version of MOSAIC as a function of the candidate number of replies used. We observe that initially, the fraction of advertisers mentioned is very high, but drops significantly once the mechanism has converged. The reason is precisely the fact that mentioning more advertisers in this case would cause either a significant drop in advertiser utility, or a drop in the usefulness of the reply to the user. By placing more weight towards the advertisers (i.e., decreasing the value of τ in Equation (1)), the platform can increase the number of advertisers shown when the mechanism has converged.

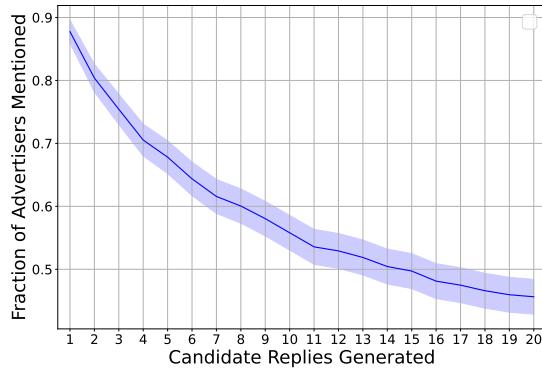


Figure 12: Fraction of advertisers mentioned in the output reply as a function of the number of candidate replies.

C.8 TESTING WITH ADDITIONAL LANGUAGE MODELS

In this section, we test MOSAIC’s performance using Google’s flan-t5-large model (Chung et al., 2022).

We repeat exactly the same procedure as detailed in Section 6.1, but using flan-t5-large as the basis for both π_{gen} , the advertisers’ LLMs, and the reference LLM. We use the same 50 user queries, each with two advertisers, and test each query on 10 different random seeds, resulting in 500 instances.

In Figure 13 to Figure 16, we present the total advertiser value gain, reward gain from participating in MOSAIC, platform revenue, and the usefulness of the returned reply to users as functions of the number of candidate replies generated.

Consistent with the main paper’s results using Llama-2-7b-chat-hf (Touvron et al., 2023), the context-aware version of MOSAIC demonstrates significant value and utility generation for advertisers and successfully captures a substantial portion as platform revenue. In contrast, though, the naive version of MOSAIC fails to generate value or utility for advertisers, resulting in negligible revenue for the platform. Similar to the results of the main paper, despite both mechanisms converging in the limit to the platform’s objective, as shown in Corollary 4.1, their performance for a computationally tractable number of generated replies diverges.

Similar to the findings in the main paper, the naive version retains the reference LLM’s usefulness to the user (which is optimal for the task), while the context-aware version prioritizes the platform’s objective, leading to lower user usefulness. Overall, our results using flan-t5-large agree closely with those obtained using Llama-2-7b-chat-hf, demonstrating consistency across models. The main difference is that the context-aware mechanism is even more important when using flan-t5-large.

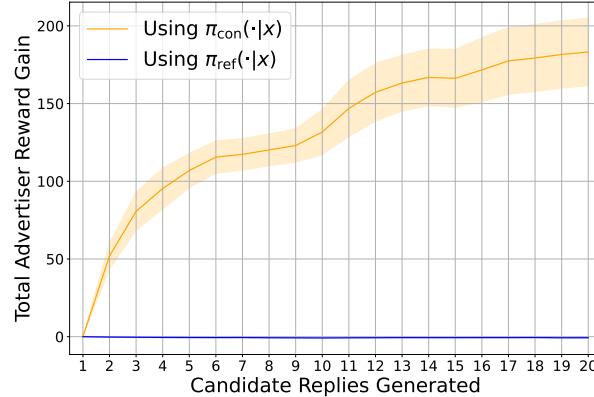


Figure 13: Advertiser value gain as a function of the number of candidate replies generated using π_{ref} and π_{con} (flan-t5-large model).

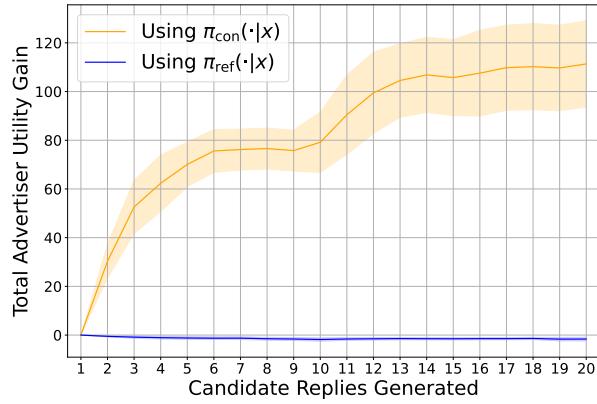


Figure 14: Advertiser value gain as a function of the number of candidate replies generated using π_{ref} and π_{con} (flan-t5-large model).

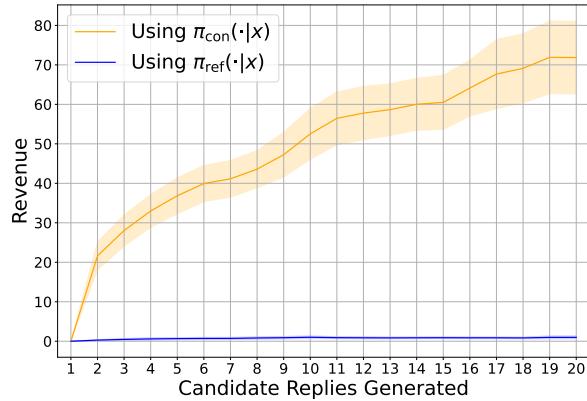


Figure 15: Revenue as a function of the number of candidate replies generated using π_{ref} and π_{con} (flan-t5-large model).

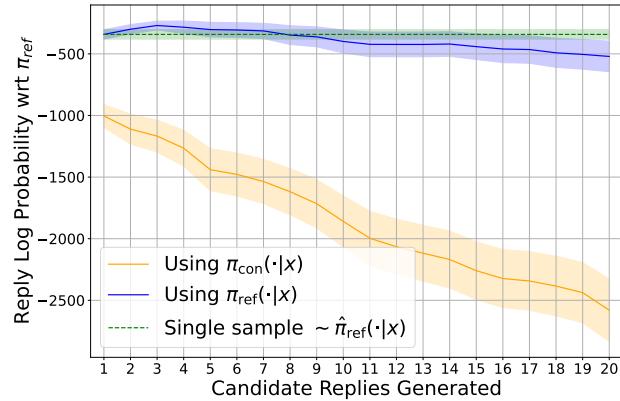


Figure 16: Reply log probability with respect to the reference LLM as a function of the number of candidate replies generated using π_{ref} and π_{con} (flan-t5-large model).

D FULL TEXT OF PROMPTS

Table 3: Table of user prompts with corresponding advertisers and their descriptions. Only the first two advertisers are shown for each query; the full prompts are available in the code supplement.

User Query 1

Learning to play the guitar.

Advertisers

- StrumMaster: selling guitars, tuners, and learning materials for beginners
 - Chordify: offering online guitar lessons and interactive practice tools
-

User Query 2

Best skin care routine for dry skin.

Advertisers

- HydraSkin: providing moisturizers and serums for dry skin
 - GlowEssentials: selling natural and organic skin care products
-

User Query 3

Improving productivity while working from home.

Advertisers

- SpaceMakers: offering ergonomic furniture and home office setups
 - FocusApp: providing productivity apps and time management tools
-

User Query 4

Starting a small online business.

Advertisers

- EcomLaunch: offering e-commerce platform solutions and web design services
 - MarketMover: providing digital marketing services and SEO optimization
-

User Query 5

Healthy meal planning on a budget.

Advertisers

- BudgetBites: selling affordable meal kits and recipe books
 - NutriSaver: offering discounts on healthy groceries and food delivery services
-

User Query 6

Mastering digital photography.

Advertisers

- PixelPro: selling cameras, lenses, and photography accessories
 - EditCraft: offering photo editing software and online tutorials
-

User Query 7

Effective ways to reduce household energy use.

Advertisers

- EcoSave: offering energy-efficient home appliances and lighting solutions
 - InsulaTech: providing home insulation and energy audit services
-

User Query 8

Finding the perfect hiking trails.

Advertisers

- TrailFinder: offering a mobile app with detailed maps and trail reviews
 - GearUp: selling outdoor gear and apparel for hiking enthusiasts
-

User Query 9

Building a personal brand on social media.

Advertisers

- BrandBuilder: offering personal branding courses and social media strategy consultations
 - VisualizeMe: providing graphic design services for social media content
-

User Query 10

Learning a new language effectively.

Advertisers

- LingoLeap: offering online language learning courses and tutoring

-
- SpeakEasy: providing language learning apps with speech recognition technology
-

User Query 11

Staying fit without a gym.

Advertisers

- HomeFit: selling home workout equipment and fitness accessories
 - MoveIt: offering online fitness classes and personal training sessions
-

User Query 12

Eco-friendly travel options.

Advertisers

- GreenPath: offering eco-friendly travel packages and sustainable tourism experiences
 - EcoStay: providing listings for green hotels and accommodations
-

User Query 13

Mastering the art of cooking steak.

Advertisers

- GrillMaster: selling premium grills and barbecue accessories
 - SteakPerfection: offering online cooking classes focused on meat preparation
-

User Query 14

Creating a successful YouTube channel.

Advertisers

- VidGrowth: offering video production courses and YouTube growth strategies
 - ChannelDesign: providing custom YouTube channel art and video thumbnails
-

User Query 15

Decorating your home on a budget.

Advertisers

- DecorDeals: selling affordable home decor and furniture
 - StyleSavvy: offering interior design consultations and budget-friendly decorating tips
-

User Query 16

Managing stress and anxiety.

Advertisers

- CalmSpace: offering mindfulness apps and stress reduction tools
 - WellnessWave: providing online therapy sessions and mental health resources
-

User Query 17

How can I make my home more energy efficient?

Advertisers

- EcoWise: offering smart home energy solutions and solar panels
 - GreenInsulate: providing eco-friendly insulation materials and installation services
-

User Query 18

What are the best strategies for investing in the stock market?

Advertisers

- InvestSmart: offering investment advice and financial planning services
 - StockGenius: providing an AI-powered stock market analysis tool
-

User Query 19

How can I improve my mental health?

Advertisers

- MindBalance: offering mental health apps and mindfulness courses
 - SerenityNow: providing online counseling and therapy sessions
-

User Query 20

Where can I find organic and sustainable clothing?

Advertisers

- EcoWear: selling sustainable and organic clothing options
 - GreenThreads: offering eco-friendly fabrics and fair-trade fashion
-

User Query 21

How do I start a vegetable garden in my backyard?

Advertisers

- GardenGuru: selling gardening tools, seeds, and starter kits
 - PlantPerfect: offering online gardening courses and personalized garden planning
-

User Query 22

What are the best tools for remote team collaboration?

Advertisers

- TeamSync: offering collaboration software and project management tools
 - ConnectSuite: providing video conferencing solutions and team communication platforms
-

User Query 23

How can I learn to cook healthy meals quickly?

Advertisers

- QuickChef: offering quick and healthy cooking classes online
 - NutriFast: selling kitchen gadgets and appliances for healthy cooking
-

User Query 24

What are the best ways to save for retirement?

Advertisers

- FutureFund: offering retirement planning and investment services
 - PensionPlus: providing personalized financial advice for retirement savings
-

User Query 25

How can I decorate my apartment on a minimalist budget?

Advertisers

- MinimalSpace: selling minimalist furniture and decor
 - SimplicityDesigns: offering interior design consultation for minimalist living spaces
-

User Query 26

What are the best practices for data security in small businesses?

Advertisers

- SecureIT: offering cybersecurity solutions for small businesses
 - DataGuard: providing data protection services and training
-

User Query 27

How can I reduce my carbon footprint with daily habits?

Advertisers

- EcoHabits: offering products and tips for sustainable living
 - GreenStep: providing carbon offset subscriptions and eco-friendly lifestyle goods
-

User Query 28

What are effective ways to learn a musical instrument online?

Advertisers

- MusicMastery: offering online music lessons and virtual tutoring
 - InstaTune: selling musical instruments and learning aids
-

User Query 29

How can I keep my pets healthy and happy?

Advertisers

- PetWell: offering pet health insurance and wellness products
 - HappyPaws: selling organic pet food and eco-friendly toys
-

User Query 30

What's the best way to start learning coding from scratch?

Advertisers

- CodeBegin: offering coding bootcamps and online tutorials for beginners
 - DevPath: providing interactive coding courses and learning platforms
-

User Query 31

How to find affordable eco-friendly travel destinations?

Advertisers

- GreenTraveler: offering eco-friendly travel guides and affordable green vacation packages
 - EcoTripAdvisor: providing reviews and bookings for sustainable travel options
-

User Query 32

What are the best ways to improve indoor air quality?

Advertisers

- AirPure: selling air purifiers and natural air freshening solutions
 - BreatheEasy: offering indoor plant subscriptions and ventilation solutions
-

User Query 33

How to make my own natural beauty products?

Advertisers

- NatureGlow: offering DIY beauty product kits and organic ingredients
 - EcoBeauty: selling guides and supplies for creating homemade natural cosmetics
-

User Query 34

What are the best techniques for effective online teaching?

Advertisers

- EduTech: offering online teaching tools and educational technology solutions
 - TeachMaster: providing courses on online pedagogy and digital classroom management
-

User Query 35

How to create a successful podcast from home?

Advertisers

- PodCastify: offering podcast creation courses and audio equipment
 - SoundWave: selling podcast editing software and hosting services
-

User Query 36

What are the most effective ways to market a small business locally?

Advertisers

- LocalBoost: offering local SEO services and marketing strategies for small businesses
 - MarketTown: providing targeted advertising solutions and local business networking opportunities
-

User Query 37

How can I optimize my home office for productivity?

Advertisers

- WorkZen: offering ergonomic office furniture and productivity tools
 - SpaceCraft: providing home office design services and organization solutions
-

User Query 38

What are the best platforms for freelance writers to find work?

Advertisers

- WriteFreely: offering a job board and resources for freelance writers
 - ContentConnect: providing a platform to connect writers with businesses
-

User Query 39

How to get started with meditation and mindfulness?

Advertisers

- MindfulStart: offering guided meditation apps and mindfulness training
 - ZenSpace: selling meditation cushions and mindfulness journals
-

User Query 40

Where can I find unique handmade gifts?

Advertisers

- CraftCorner: selling handmade gifts and artisan crafts
 - MakerMarket: offering a marketplace for unique handmade items from local artists
-

User Query 41

How can I improve my credit score quickly?

Advertisers

- CreditBoost: offering credit monitoring services and financial advice
 - ScoreSurge: providing tools and tips for improving your credit score
-

User Query 42

What are the best apps for tracking fitness and nutrition?

Advertisers

- FitTrack: offering comprehensive fitness and nutrition tracking apps

-
- NutriSync: selling personalized nutrition plans and diet tracking tools
-

User Query 43

How to plan an environmentally friendly wedding?

Advertisers

- GreenBride: offering eco-friendly wedding planning services and supplies
 - EcoWed: providing sustainable wedding attire and decor
-

User Query 44

What are the safest ways to travel during a pandemic?

Advertisers

- SafeJourney: offering travel safety kits and pandemic travel advice
 - HealthTravel: providing information on safe destinations and travel insurance
-

User Query 45

How can I learn DIY home repairs and improvements?

Advertisers

- FixItYourself: offering online courses and tutorials for home repair
 - DIYHome: selling DIY home improvement tools and kits
-

User Query 46

What are the best resources for starting a plant-based diet?

Advertisers

- PlantEats: offering guides and meal plans for starting a plant-based diet
 - VeggieVibe: selling plant-based cookbooks and kitchen gadgets
-

User Query 47

Tips for running a marathon.

Advertisers

- RunFastGear: specializing in high-performance running shoes
 - HydraFuel: offering electrolyte drinks and energy bars for athletes
-

User Query 48

Best practices for organic gardening.

Advertisers

- GreenThumb Solutions: providing organic fertilizers and pest control products
 - EcoGrow: selling heirloom seeds and sustainable gardening tools
-

User Query 49

How to improve home WiFi?

Advertisers

- SignalBoost: which offers advanced WiFi routers and extenders
 - NetWizard: providing network optimization services and support
-

User Query 50

Ways to save on travel.

Advertisers

- BudgetJourneys: specializing in affordable travel packages and deals
 - StayLocal: offering discounts on boutique hotels and unique accommodations
-