

Social Welfare Maximization in Approval-Based Committee Voting under Uncertainty

Haris Aziz
UNSW Sydney
Sydney, Australia
haris.aziz@unsw.edu.au

Yuhang Guo
UNSW Sydney
Sydney, Australia
yuhang.guo2@unsw.edu.au

Venkateswara Rao Kagita
National Institute of Technology
Warangal, India
venkat.kagita@nitw.ac.in

Baharak Rastegari
University of Southampton
Southampton, United Kingdom
b.rastegari@soton.ac.uk

Mashbat Suzuki
UNSW Sydney
Sydney, Australia
mashbat.suzuki@unsw.edu.au

ABSTRACT

Approval voting is widely used for making multi-winner voting decisions. The canonical rule (also called Approval Voting) used in the setting aims to maximize social welfare by selecting candidates with the highest number of approvals. We revisit approval-based multi-winner voting in scenarios where the information regarding the voters' preferences is uncertain. We present several algorithmic results for problems related to social welfare maximization under uncertainty, including computing the social welfare probability distribution of a given outcome, computing the probability that a given outcome is social welfare maximizing, computing an outcome that is social welfare maximizing with the highest probability, and understanding how robust an outcome is with respect to social welfare maximization.

KEYWORDS

Approval-Based Committee Voting; Uncertain Preference; Social Welfare Maximization

ACM Reference Format:

Haris Aziz, Yuhang Guo, Venkateswara Rao Kagita, Baharak Rastegari, and Mashbat Suzuki. 2026. Social Welfare Maximization in Approval-Based Committee Voting under Uncertainty. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, Paphos, Cyprus, May 25 – 29, 2026, IFAAMAS, 9 pages. <https://doi.org/10.65109/RABQ9914>

1 INTRODUCTION

Approval voting is one of the simplest and most widely used methods of making selection decisions. Due to its fundamental nature, it has found applications in recommender system [14, 17, 26], blockchains [9, 13], and Q & A platforms [21]. In approval voting, voters are asked to identify the candidates they approve of from a given set. The candidates with the highest number of approvals are then selected. Therefore, “approval voting” not only specifies the format of the ballots but also commonly points to the method for selecting the candidates [22, 23]. Many organizations and societies

use approval voting to select committees. For example, the Institute of Electrical and Electronics Engineers (IEEE), one of the largest scientific and technical organizations, has been using approval voting for selection decisions.

If approvals of voters are interpreted as voters' binary preferences over candidates, then the outcome of the approval voting method has a clear *utilitarian social welfare* perspective: identify the set of candidates that provide the highest social welfare to the voters. We explore this utilitarian social welfare perspective when there is uncertain information regarding voters' preferences. Uncertain approval preferences are useful when the central planner only has imprecise information about the voters' preferences. This estimated information could be based on historical preferences, past selections, or online clicks or views. For example, if an agent i has selected a certain candidate c 70% of the time in previous situations, one could use this information to assume that the approval probability of agent i for candidate c is 0.7. The uncertain information could also be based on situations where each agent represents a group of people who may not have identical approval preferences. For example, if 60% of the group represented by agent i approved a certain candidate c , one could assume that the approval probability of agent i for candidate c is 0.6. Uncertainty becomes a prevalent factor also when employing methods such as machine learning or recommendation techniques to forecast the unobserved (dis)approvals of voters for candidates. (The motivating examples are by Aziz et al. [6].)

We consider four different types of uncertain approval preferences that have been studied in recent work (see, e.g., [6]). Under the *Candidate Probability model*, there is a probability for a given voter approving a given candidate. This model captures the examples in the previous paragraph. The Three Valued Approval (3VA) model [20] is a restricted version of the candidate probability model and captures a natural form of uncertainty where a voter has no or too little information on some candidates and assigns approval probability of 0.5 to such candidates. Under the *Lottery model*, each voter has an independent probability distribution over approval sets. Under the *Joint Probability model*, there is a probability distribution of approval profiles. These last two models allow us to capture richer forms of uncertainty where there could be dependencies between candidates and between voters' approval sets. The Joint Probability model, in particular, may not seem practical; we include it for completeness.



This work is licensed under a Creative Commons Attribution International 4.0 License.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). <https://doi.org/10.65109/RABQ9914>

In classical approval-based committee voting, the social welfare of any given committee can be computed exactly, making it straightforward to evaluate how “good” the committee is. However, in scenarios involving uncertain preferences, the social welfare of the given committee becomes a random variable and a natural, fundamental question arises: What is the probability that a given committee achieves a sufficiently high level of social welfare? More generally, what is the distribution of the social welfare associated with the committee? Understanding the distribution is essential, as it captures the likelihood of the given committee achieving desirable social welfare. With social welfare as the objective, the optimal committee is naturally the one that maximizes social welfare. Under deterministic preferences, this problem is well-understood. For any given plausible approval profile, the Approval Voting (AV) rule efficiently identifies the committee that maximizes social welfare in polynomial time. However, in scenarios under uncertain preference models, for example, the Lottery model or Candidate Probability model, a desirable committee in this context is not just one that performs well in a single realization, but one that maximizes social welfare in the largest fraction of all possible realizations, that is, the committee with the highest probability of being social welfare maximizing. When facing uncertainty, another intriguing and practically relevant question concerns the robustness of a committee. For example, under uncertain preferences, does there exist a committee that achieves at least half of the optimal social welfare with high probability? This shift from deterministic to probabilistic evaluation introduces significant computational challenges. Since the number of plausible approval profiles grows exponentially with the number of agents and candidates, problems related to social welfare maximizing committees are computationally challenging. This leads to compelling questions:

How can we compute the distribution of a given committee’s social welfare? Can we compute the probability of a given committee being social welfare maximizing in polynomial time? Can we efficiently identify the committee that maximizes social welfare (or well-approximates optimal social welfare) with the highest probability?

1.1 Our Results

Our first contribution is polynomial-time algorithms for computing the distribution of a given committee’s social welfare (SW-DIST). We design algorithms both for Lottery and Candidate probability models. These results allow us to check whether a committee achieves high level of social welfare with high probability.

We next turn to the problem of computing the probability that a given committee is social welfare maximizing (SWM-PROB(W, p)). We show that under the Candidate Probability model, the problem is solvable in polynomial time. In contrast, we prove that the problem is NP-complete under the Lottery model. We also present a positive result for the decision variant that asks whether a given committee is social welfare maximizing with probability one, showing that it is solvable in polynomial time.

Next, we explore the problem whether there exists a committee whose probability of being social welfare maximizing exceeds a

given threshold p (EXISTS-SWM-PROB(p)). We show that this problem is NP-hard under the Lottery model. Nevertheless, we provide a polynomial-time result for the special case where $p = 1$. For the Candidate Probability model, the problem becomes intriguing as even for the more restrictive 3-Valued Approval (3VA) setting, there is no known characterization of committees that maximize the probability of being social welfare maximizing. While the complexity of the problem remains open, we present positive results under certain constraints, for instance, when the committee size k is constant, or when $p = 1$.

We finally consider the problem of robust welfare maximization: does there exist a committee which guarantees a fraction of the optimal social welfare of the realized profile with high probability? We provide a positive result for the 3VA model and an impossibility result for the Candidate Probability model. Our key results are summarized in Table 1. Missing proofs are relegated to the appendix.

Problems	Lottery Model	Candidate Probability Model
SW-DIST	in P (Thm. 3.1)	in P (Thm. 3.2)
SWM-PROB(W, p)	NP-h (Thm. 4.1) (in P (Thm. 4.4) when $p = 1$)	in P (Thm. 4.5)
EXISTS-SWM-PROB(p)	NP-h (Thm. 5.1) (in P (Thm. 5.3) when $p = 1$)	?

Table 1: Summary of results.

1.2 Related Work

Approval-Based Committee (ABC) voting has received considerable attention in recent years (see, e.g., [3, 4, 12, 23, 25]), primarily focusing on selecting “proportional” committees. Concurrently, there has been growing interest in preference aggregation under uncertainty. Konczak and Lang [24] study winner determination problems with incomplete preferences for ranking-based single-winner voting rules. Hazon et al. [19] explore the probability of a particular candidate winning an election under uncertain preferences for various voting rules, such as the Plurality rule and the Borda rule. Boutilier and Rosenschein [10] provide a survey on uncertainty and communication in voting. Do et al. [15] investigate the dynamic selection of candidates, where uncertainty is related to the order in which candidates appear. Halpern et al. [18] devise different query algorithms for scenarios where voters’ ballots are partial and incomplete over all candidates, while Brill et al. [11] propose an ABC voting model with possibly unavailable candidates and examine voting rules which admit “safe” query policies to check candidates’ availability.

Highly relevant to our paper, Barrot et al. [8] examine single and multi-winner voting under uncertain approvals, and Terzopoulou et al. [27] address the problem of checking whether an incomplete approval profile admits a completion within a certain restricted domain of approval preferences. Imber et al. [20] study several computational problems, including checking whether a given committee is possibly or necessarily *Justified Representation* (JR) or whether there is a possible outcome of various rules, including Approval Voting. Of the uncertain preference models that we consider, Imber et al. [20] have explored the 3VA model. Most of the computational

problems that we consider are not studied by Imber et al. [20]. They presented a polynomial-time algorithm to check whether given a committee is social welfare maximizing under some realization of the uncertain preferences. This result is implied by our more general results. Recently, Aziz et al. [6] consider several probabilistic preference models and problems, such as maximizing the probability of satisfying JR. In our paper, we consider the fundamental objective of maximizing social welfare under the same uncertain preference models.

Uncertain preferences have also been studied in other domains, such as matching and fair allocations. Aziz et al. [1] investigate the computational complexity of Pareto optimal allocation under uncertain preferences. Aziz et al. [5] explore the envy-free allocation for the house allocation market. Aziz et al. [2] address the problem of computing stable matchings with uncertain preferences. Additionally, Bampis et al. [7] examine stable matchings under one-sided uncertainty, focusing on computational issues within three different competitive query models.

2 PRELIMINARIES

For any $t \in \mathbb{N}$, let $[t] := \{1, 2, \dots, t\}$. An *instance* of the (deterministic) approval-based committee (ABC) voting is represented as a tuple (V, C, \mathcal{A}, k) , where:

- $V = [n]$ and $C = [m]$ are the sets of *voters* and *candidates*, respectively.
- Each voter approves a set of candidates $A_i \subseteq C$. Let $\mathcal{A} = (A_1, A_2, \dots, A_n)$ denote voters' approval profile. The set of all possible approval profiles is denoted by \mathbf{A} .
- k is a positive integer that represents the committee size.

A *feasible winning committee* $W \subseteq C$ is of size k . In this paper, we consider only feasible winning committees unless otherwise specified. Given a committee W , the welfare of each voter i is the number of candidates in W of whom i approves. The *Social Welfare* (SW) of W given the approval profile \mathcal{A} is defined as the sum of the welfare of the voters, $\text{SW}(W, \mathcal{A}) = \sum_{i \in V} |W \cap A_i|$. Given an approval profile $\mathcal{A} = (A_1, A_2, \dots, A_n)$, for each candidate $c \in C$, we denote the approval score of candidate c by $\text{AS}(c, \mathcal{A}) = |\{i \in V : c \in A_i\}|$. A committee W is *Social Welfare Maximizing* (SWM) under approval profile \mathcal{A} if W generates the maximum social welfare among all committees of size $|W| = k$. Given an approval profile \mathcal{A} , we can compute an SWM committee in polynomial time by computing each candidate's approval score and selecting the k candidates with the highest approval scores in a greedy manner. Consequently, given a deterministic approval profile \mathcal{A} and a committee W , we can decide in polynomial time whether or not W is SWM.

2.1 Uncertain Preference Models

Our main focus is on ABC voting under *uncertain* approval ballots, where there is an underlying probability distribution over approval profiles. We adopt the following uncertainty models considered by Aziz et al. [6].

- (1) **Joint Probability model:** A probability distribution $\Delta(\mathbf{A}) := \{(\lambda_r, \mathcal{A}_r)\}_{r \in [s]}$ is given over s possible approval profiles with $\sum_{r \in [s]} \lambda_r = 1$, where for each $r \in [s]$, the approval profile \mathcal{A}_r is associated with a positive probability $\lambda_r > 0$. We write $\Delta(\mathcal{A}_r) = \lambda_r$.

- (2) **Lottery model:** For each voter $i \in V$, we are given a probability distribution $\Delta_i := \{(\lambda_r, S_r)\}_{r \in [s_i]}$ over s_i approval sets with $\sum_{r \in [s_i]} \lambda_r = 1$. For each $r \in [s_i]$, voter i approves (exactly) candidate set $S_r \subseteq C$ with probability $\lambda_r > 0$. We write $\Delta_i(S_r) = \lambda_r$. We assume that the probability distributions of voters are independent.
- (3) **Candidate Probability model:** Each voter i approves each candidate c *independently* with probability $p_{i,c}$, i.e., for each $i \in V$ and each $c \in C$, $p_{i,c} \in [0, 1]$. The **Three Valued Approval (3VA) model**, is a special case where each agent specifies a subset of candidates that are approved and a subset of candidates that are disapproved, and the remaining candidates could be approved or disapproved independently with equal probability. Specifically, $\forall i \in V, c \in C, p_{i,c} \in \{0, \frac{1}{2}, 1\}$, wherein 0 denotes disapproval, 1 indicates approval, and $\frac{1}{2}$ represents unknown.

We use the notation $\Delta := (\Delta_1, \Delta_2, \dots, \Delta_n)$ to represent the input of the uncertain approval model, which may follow either the Lottery model or the Candidate Probability model, depending on the context of the problem or algorithm under consideration. Under the Lottery model, each $\Delta_i := (\lambda_r, S_r)_{r \in [s_i]}$ specifies a probability distribution over s_i approval sets for agent i . For the Candidate Probability model, each element $\Delta_i := (p_{i,1}, p_{i,2}, \dots, p_{i,m})$ represents agent i 's independent approval probabilities over the m candidates.

The Joint Probability and Lottery models have been studied in other contexts including two-sided stable matching problems and assignment problems [1, 2]. The 3VA model has been studied in ABC voting [20]. We refer to an approval profile that occurs with positive probability, under any of the uncertainty models, as a *plausible* profile. The following facts about these uncertainty models were recently pointed out by Aziz et al. [6].

PROPOSITION 2.1 (AZIZ ET AL. [6]). *There is a unique Joint Probability model representation for preferences given under the Lottery model; There is a unique Lottery model representation for preferences given in Candidate-Probability model.*

2.2 Computational Problems

We begin by considering the most natural MAXEXP SW problem, which computes a committee that maximizes the expected social welfare ($\mathbb{E}[\text{SW}(W)]$). Formally, $\mathbb{E}[\text{SW}(W)]$ of a committee W is defined as $\mathbb{E}[\text{SW}(W)] = \sum_{\mathcal{A} \in \mathbf{A}} \Delta(\mathcal{A}) \cdot \text{SW}(W, \mathcal{A}) = \sum_{\mathcal{A} \in \mathbf{A}} \Delta(\mathcal{A}) \cdot \sum_{i \in V} |W \cap A_i|$. We show that it is polynomial-time solvable¹ under any of the four studied uncertain models and defer the proofs to the appendix. After this, we study a fundamental computational problem: given a committee W , how can we compute the distribution of its social welfare under an uncertain approval model? We term this problem SW-DIST. Notice that for any fixed committee W , its social welfare is a discrete random variable that takes integer values in the range $[0, kn]$. Formally, the SW-DIST problem entails computing the probability $\Pr[\text{SW}(W) = \tau]$ for each $\tau \in [0, kn]$.

We next consider the problem of computing the exact probability that a given committee W is social welfare maximizing (SWM) under uncertain preferences. We formalize this as a decision problem termed SWM-PROB(W, p), which is defined as follows: given a

¹Here and throughout, "polynomial-time solvable" means polynomial in the size of the input.

committee W and a threshold $p \in [0, 1]$, determine whether W has at least probability p of being social welfare maximizing.

SWM-PROB(W, p)

Input: Voters V , Uncertain Approval Model Δ , Committee W , threshold p ;

Question: Decide whether $\Pr[W \text{ is SWM}] \geq p$.

Beyond computing the probability for a fixed committee being social welfare maximizing, a more challenging problem we investigate in this paper is how to identify a committee that has the highest probability of being social welfare maximizing. We formally formulate this as a decision problem, denoted by EXISTS-SWM-PROB(p): given an uncertain approval model and a threshold $p \in [0, 1]$, decide whether there exists a committee W such that the probability of W being social welfare maximizing is at least p .

EXISTS-SWM-PROB(p)

Input: Voters V , Uncertain Approval Model Δ , threshold p ;

Question: Decide whether there exists a committee W such that $\Pr[W \text{ is SWM}] \geq p$.

The rest of the paper is organized as follows. In Section 3 we study the SW-DIST problem under the Lottery and Candidate Probability models and propose two distinct dynamic programming algorithms. With this fundamental tool in hand, we examine the SWM-PROB(W, p) problem in Section 4 and provide hardness results under the Lottery model and polynomial-time results under the Candidate Probability model. In Section 5, we study the EXISTS-SWM-PROB(p) problem. Finally, we define robust committees and discuss the existence and computation of a robust committee under the Candidate Probability model in Section 6.

3 COMPUTATION OF SOCIAL WELFARE DISTRIBUTION

We begin by examining the fundamental SW-DIST problem: Given a committee W , what is the distribution of its social welfare under uncertain preference models? As previously discussed, the social welfare $SW(W)$ is a discrete random variable that takes integer values in the range $[0, kn]$. Hence, we can focus on the key computational problem: given a committee W and an integer $\tau \in [0, kn]$, compute the probability $\Pr[SW(W) = \tau]$. By solving this for all values of τ in the range $[0, kn]$, we can fully characterize the distribution of $SW(W)$. Under the Lottery and Candidate Probability models, we propose two distinct dynamic programming algorithms, each solving the SW-DIST problem in polynomial time.

THEOREM 3.1. *Under the Lottery model, SW-DIST is solvable in polynomial time.*

PROOF. Given a committee W and a deterministic approval profile \mathcal{A} , $SW(W)$ is the sum of $|W \cap A_i|$ for each voter i . Under the Lottery model, each voter's approval set is random and hence the value of $|W \cap A_i|$ is a random variable, denoted by $f_i(W)$. Then,

we have $\Pr[SW(W) = \tau] = \Pr[\sum_{i \in [n]} f_i(W) = \tau]$. Conditioning on $f_n(W)$, it can be reformulated as $\sum_{r=0}^{\tau} \Pr[\sum_{i \in [n-1]} f_i(W) = r \mid f_n(W) = (\tau - r)] \cdot \Pr[f_n(W) = (\tau - r)]$. Note that for any two voters i and j , $f_i(W)$ is independent of $f_j(W)$ as each voter's approval set is sampled independently. Therefore

$$\begin{aligned} \Pr[SW(W) = \tau] &= \Pr\left[\sum_{i \in [n]} f_i(W) = \tau\right] \\ &= \sum_{r=0}^{\tau} \Pr\left[\sum_{i \in [n-1]} f_i(W) = r\right] \cdot \Pr[f_n(W) = (\tau - r)]. \end{aligned} \quad (1)$$

We design Algorithm 1 by leveraging Equation (1). In Algorithm 1,

Algorithm 1 SW-DIST algorithm for the Lottery model

Input: W, k, τ, Δ

Output: $\text{dp}[n][\tau]$

- 1: Initialize an $n \cdot (k + 1)$ matrix f with elements 0;
 - 2: **for** each voter $i \in V$ **do**
 - 3: **for** integer $j = (0, \dots, k)$ **do**
 - 4: $f[i][j] \leftarrow \sum_{r \in [s_i]} \lambda_r \cdot \mathbb{I}[|W \cap S_r| = j]$;
 - 5: **end for**
 - 6: **end for**
 - 7: Initialize an $n \cdot (nk + 1)$ matrix dp with elements 0;
 - 8: **for** $t = (0, \dots, k)$, $\text{dp}[1][t] \leftarrow f[1][t]$;
 - 9: **for** $i \leftarrow (2, \dots, n)$ **do**
 - 10: **for** $t \leftarrow (0, \dots, nk)$ **do**
 - 11: $\text{dp}[i][t] = \sum_{r=0}^t \text{dp}[i-1][r] \cdot f[i][t-r]$;
 - 12: **end for**
 - 13: **end for**
 - 14: Return $\text{dp}[n][\tau]$.
-

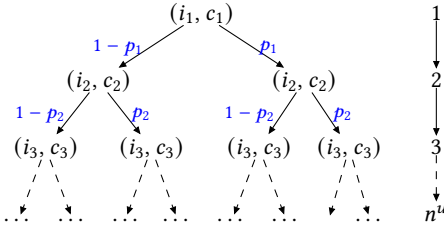
we denote $f[i][j]$ as the probability of the event $f_i(W) = j$ (line 2-6). Specifically, $f[i][j]$ represents the sum of the realization probabilities λ_r of the deterministic approval sets S_r where the size of the intersection with W is exactly j . After pre-processing, we initialize the dynamic programming matrix dp , where $\text{dp}[i][j]$ is the probability of $\sum_{\ell \in [i]} f_\ell(W) = j$. Note that, computing $\Pr[SW(W) = \tau]$ is equivalent to determining the value of $\text{dp}[n][\tau]$. The recursive relation in Equation (1) corresponds to $\text{dp}[n][\tau] = \sum_{r=0}^{\tau} \text{dp}[n-1][r] \cdot f[n][\tau-r]$. Starting from $\text{dp}[1][0]$, we compute each value in the dp matrix recursively (line 8-13). Since the algorithm runs in $O(nk \cdot \max(k, \max_{i \in N} |s_i|))$ time, it implies that SW-DIST under the Lottery model is solvable in polynomial time. \square

THEOREM 3.2. *Under the Candidate Probability model, SW-DIST is solvable in polynomial time.*

PROOF. Under the candidate probability model, given a committee W , for each voter i and each candidate $c \in W$, $p_{i,c}$ falls into three different cases:

- $p_{i,c} = 0$, voter i certainly disapproves candidate c ;
- $p_{i,c} = 1$, voter i certainly approves candidate c ;
- $p_{i,c} \in (0, 1)$, voter i approves candidate c with an uncertain probability $p_{i,c}$.

We first denote $n^1 = |\{(i, c) : i \in V, c \in W, p_{i,c} = 1\}|$ as the number of certain approvals while $n^u = |\{(i, c) : i \in V, c \in W, p_{i,c} \in (0, 1)\}|$ as the number of uncertain approvals. Because of the existence of the uncertain approvals, the social welfare $\text{SW}(W)$ of the given committee W is a random variable ranging from n^1 to $n^1 + n^u$. Furthermore, $\text{SW}(W)$ is distributed according to shifted Poisson binomial distribution with n^u independent Bernoulli trials. We first relabel these uncertain approval pairs as $(i_1, c_1), (i_2, c_2), \dots, (i_{n^u}, c_{n^u})$ and the corresponding success probabilities as $(p_1, p_2, \dots, p_{n^u})$. Intuitively, we represent all the realization of these uncertain approval pairs as a tree as follows.



Every path in the tree represents a specific realization of uncertain approvals (trials) transforming into approvals (success) or disapprovals (failure). For SW-DIST problem $\Pr[\text{SW}(W) = \tau]$, if $\tau < n^1$ or $\tau > n^1 + n^u$, $\Pr[\text{SW}(W) = \tau] = 0$. So we mainly focus on $\tau \in [n^1, n^1 + n^u]$. Denote $t = \tau - n^1$. Then $\Pr[\text{SW}(W) = \tau]$ can be represented as follows.

$$\begin{aligned} \Pr[\text{SW}(W) = \tau] &= \Pr[\text{SW}(W) - n^1 = t] \\ &= \Pr[t \text{ out of } n^u \text{ trials succeed}] \\ &= \left(\Pr[(t-1) \text{ succeed in } (n^u-1) \text{ trials}] \cdot \Pr[n^u\text{-th trial succeeds}] \right) \\ &\quad + \left(\Pr[t \text{ out of } (n^u-1) \text{ trials succeed}] \cdot \Pr[n^u\text{-th trial fails}] \right). \end{aligned} \quad (2)$$

Based on the above Equation (2), we provide the following dynamic programming Algorithm 2 to solve SW-DIST problem. As we

Algorithm 2 SW-DIST algorithm for the Candidate Probability model

Input: W, n^u, t, Δ .

Output: $\text{dp}[n^u][t]$.

- 1: Initialize an $n^u \cdot (t+1)$ matrix dp with elements 0;
 - 2: $\text{dp}[1][0] \leftarrow (1-p_1), \text{dp}[1][1] \leftarrow p_1$;
 - 3: **for** $i \leftarrow (2, \dots, n^u)$ **do**
 - 4: **for** $j \leftarrow (1, \dots, t)$ **do**
 - 5: $\text{dp}[i][j] \leftarrow (p_i \cdot \text{dp}[i-1][j-1]) + ((1-p_i) \cdot \text{dp}[i-1][j])$;
 - 6: **end for**
 - 7: **end for**
 - 8: Return $\text{dp}[n^u][t]$.
-

re-labeled the n^u uncertain approval pairs (independent Bernoulli trials), in Algorithm 2, $\text{dp}[i][j]$ represents the probability that there are j trials which succeed among the first i trials. Then, computing $\Pr[\text{SW}(W) = \tau]$ is equivalent to computing $\text{dp}[n^u][t]$. According to Equation (2), $\text{dp}[n^u][t] = p_{n^u} \cdot \text{dp}[n^u-1][t-1] + (1-p_{n^u}) \text{dp}[n^u-1][t]$, corresponding to line 5 in Algorithm 2. For the computation,

we first initialize $\text{dp}[1][1]$ and $\text{dp}[1][0]$, which represent the probability of success or failure of the first trial (i_1, c_1) , respectively. This is equal to the probability that voter i_1 approves (disapproves) candidate c_1 , respectively (line 2). From lines 3 to 7, Algorithm 2 recursively computes $\text{dp}[i][j]$. Finally, Algorithm 2 returns $\text{dp}[n^u][t]$, which is equal to $\Pr[\text{SW}(W) = \tau]$. We conclude that SW-DIST problem is solvable in polynomial time under the Candidate Probability model as Algorithm 2 runs in $O(mn \cdot kn) = O(kmn^2)$ time. \square

4 PROBABILITY OF A COMMITTEE BEING WELFARE MAXIMIZING

With the fundamental tools for computing the social welfare distribution in place, we now turn to the significant problem of $\text{SWM-PROB}(W, p)$, which is to decide, given a committee W and a threshold p , whether $\Pr[W \text{ is SWM}] \geq p$. While we have shown that SW-DIST can be computed in polynomial time under both Lottery and Candidate Probability models, the complexity landscape changes significantly when it comes to $\text{SWM-PROB}(W, p)$. In particular, under the Lottery model, we establish the computational intractability of the problem in general, and present intriguing differences between certain special cases. Specifically, for some parameter $\varepsilon > 0$, $\text{SWM-PROB}(W, \varepsilon)$ is NP-hard, whereas $\text{SWM-PROB}(W, 1)$ can be solved in polynomial time. In contrast, we are able to utilize the SW-DIST solution tool to show that $\text{SWM-PROB}(W, p)$ is solvable in polynomial time under the Candidate Probability model.

THEOREM 4.1. *Under the Lottery model, $\text{SWM-PROB}(W, p)$ is NP-hard.*

PROOF SKETCH. Consider the problem of checking whether it is possible for a given committee W to be SWM, i.e., deciding whether $\Pr[W \text{ is SWM}] > 0$. We first show that there is a polynomial-time reduction from this problem to $\text{SWM-PROB}(W, \varepsilon)$ where $\varepsilon < \prod_{i \in [n]} \min_{r \in [s_i]} \{\lambda_r\}$. We next prove that for any committee W , deciding whether $\Pr[W \text{ is SWM}] > 0$ is NP-hard even when $k = 1$ by reducing from the Exact Cover by 3-Sets (X3C) problem [16]. \square

Moreover, there exists a one-to-one correspondence in the reduction construction between each realization under which the given committee W is SWM and each solution associated with the X3C instance. The next corollary thus follows.

COROLLARY 4.2. *Under the Lottery model, given a committee W , computing $\Pr[W \text{ is SWM}]$ is #P-complete.*

If we set the parameter $p = 1$, that is, the $\text{SWM-PROB}(W, 1)$ problem, we obtain a polynomial-time result. The key idea is to carefully construct a deterministic approval profile and prove that any committee W is necessarily SWM if and only if W is SWM under the constructed deterministic approval profile. Before the formal proof, we introduce the following lemma.

LEMMA 4.3. *Under the Lottery model, given a committee W , it is a YES instance for $\text{SWM-PROB}(W, 1)$ if and only if, for every candidate pair (c, c') where $c \in W$ and $c' \in C \setminus W$, and for every plausible approval profile \mathcal{A} , the approval score of c is at least as large as the approval score of c' , that is, $\text{AS}(c, \mathcal{A}) \geq \text{AS}(c', \mathcal{A})$.*

With Lemma 4.3 in hand, to prove that $\text{SWM-PROB}(W, 1)$ is in P under the Lottery model, it is sufficient to show that it can be

checked in polynomial time whether, for all candidate pairs (c, c') where $c \in W$ and $c' \in C \setminus W$, the condition $\text{AS}(c, \mathcal{A}) \geq \text{AS}(c', \mathcal{A})$ holds for every plausible approval profile \mathcal{A} .

THEOREM 4.4. *Under the Lottery model, $\text{SWM-PROB}(W, 1)$ is solvable in polynomial time.*

PROOF SKETCH. For each candidate pair (c, c') , we construct a *deterministic* approval profile $\bar{\mathcal{A}}$ and demonstrate that if (c, c') satisfies $\text{AS}(c, \bar{\mathcal{A}}) \geq \text{AS}(c', \bar{\mathcal{A}})$ for the constructed profile $\bar{\mathcal{A}}$, then (c, c') satisfies $\text{AS}(c, \mathcal{A}) \geq \text{AS}(c', \mathcal{A})$ for all plausible approval profiles \mathcal{A} . The construction is as follows. Given a committee W , for each pair (c, c') where $c \in W$ and $c' \in C \setminus W$ and each voter's approval set A_i , there are four possible cases: (1) $c' \in A_i$ and $c \notin A_i$; (2) $c' \in A_i$ and $c \in A_i$; (3) $c' \notin A_i$ and $c \notin A_i$; (4) $c' \notin A_i$ and $c \in A_i$. We construct the deterministic profile $\bar{\mathcal{A}}$ as follows: for each voter i , set \bar{A}_i by selecting a plausible approval set in the following priority order: (1) \succ (2) \succ (3) \succ (4). That is, we first check whether there exists a plausible approval set such that c' is in the approval set while c is not. If such an approval set exists, we set it as \bar{A}_i in the deterministic approval profile $\bar{\mathcal{A}}$; otherwise, we consider cases (2), (3), and (4) in sequence.

Next, we prove that if a pair (c, c') satisfies $\text{AS}(c, \bar{\mathcal{A}}) \geq \text{AS}(c', \bar{\mathcal{A}})$ in the constructed profile $\bar{\mathcal{A}}$, then the pair (c, c') satisfies $\text{AS}(c, \mathcal{A}) \geq \text{AS}(c', \mathcal{A})$ for all plausible approval profiles \mathcal{A} .

Since the construction and verification can be computed in polynomial time, $\text{SWM-PROB}(W, 1)$ under the Lottery model is solvable in polynomial time. \square

We now turn into the Candidate Probability model and show that, under this model, the $\text{SWM-PROB}(W, p)$ problem is solvable in polynomial time by showing $\Pr[W \text{ is SWM}]$ is polynomial-time computable.

THEOREM 4.5. *Under the Candidate Probability model, the problem $\text{SWM-PROB}(W, p)$ is solvable in polynomial time.*

PROOF. Given a committee W , we may re-label the candidates, without loss of generality, so that $W = \{c_1, \dots, c_k\}$. We denote $\text{AS}(c)$ as a random variable corresponding to the approval score of candidate $c \in C$. Under the Candidate Probability model, the probability that a committee W is SWM is equivalent to the probability of sampling an approval profile where the approval scores of the candidates in $W = \{c_1, \dots, c_k\}$ rank among the top- k .

$$\begin{aligned} \Pr[W \text{ is SWM}] &= \sum_{\mathcal{A} \in \mathcal{A}} \Delta(\mathcal{A}) \cdot \mathbb{I}[\text{ISWMM}(W, \mathcal{A})] \\ &= \sum_{\mathcal{A} \in \mathcal{A}} \Delta(\mathcal{A}) \cdot \mathbb{I}[\text{AS}(c_1, \mathcal{A}), \dots, \text{AS}(c_k, \mathcal{A}) \text{ rank top-}k] \\ &= \Pr \left[\max_{c \in C \setminus W} \{\text{AS}(c)\} \leq \min_{1 \leq i \leq k} \{\text{AS}(c_i)\} \right]. \end{aligned}$$

Conditioning on the value of $\min_{1 \leq i \leq k} \text{AS}(c_i)$, we rewrite the probability $\Pr \left[\max_{c \in C \setminus W} \{\text{AS}(c)\} \leq \min_{1 \leq i \leq k} \{\text{AS}(c_i)\} \right]$ as

$$\sum_{t=0}^n \Pr \left[\min_{1 \leq i \leq k} \{\text{AS}(c_i)\} = t \right] \cdot \Pr \left[\max_{c \in C \setminus W} \{\text{AS}(c)\} \leq t \mid \min_{1 \leq i \leq k} \{\text{AS}(c_i)\} = t \right].$$

For any two candidates $c_i, c_j \in C$, $\text{AS}(c_i)$ is independent of $\text{AS}(c_j)$ because for each voter $v \in V$, the event that v approves c_i is independent of the event where v approves c_j . Notably, for the conditional probability $\Pr[\min_{1 \leq i \leq k} \{\text{AS}(c_i)\} = t]$, the random variable $\min_{1 \leq i \leq k} \{\text{AS}(c_i)\}$ only depends on $\{\text{AS}(c_1), \dots, \text{AS}(c_k)\}$ and is independent of $\{\text{AS}(c_{k+1}), \dots, \text{AS}(c_m)\}$. It follows that

$$\begin{aligned} &\Pr \left[\max_{c \in C \setminus W} \{\text{AS}(c)\} \leq \min_{1 \leq i \leq k} \{\text{AS}(c_i)\} \right] \\ &= \sum_{t=0}^n \Pr \left[\min_{1 \leq i \leq k} \{\text{AS}(c_i)\} = t \right] \cdot \left(\prod_{c \in C \setminus W} \Pr[\text{AS}(c) \leq t] \right). \end{aligned}$$

Now it boils down to compute $\Pr[\min_{1 \leq i \leq k} \{\text{AS}(c_i)\} = t]$, which can be reformulated as

$$\begin{aligned} &\Pr \left[\min_{1 \leq i \leq k} \{\text{AS}(c_i)\} = t \right] \\ &= \Pr \left[\min_{1 \leq i \leq k} \{\text{AS}(c_i)\} > (t-1) \right] - \Pr \left[\min_{1 \leq i \leq k} \{\text{AS}(c_i)\} > t \right] \\ &= \prod_{1 \leq i \leq k} \Pr[\text{AS}(c_i) > (t-1)] - \prod_{1 \leq i \leq k} \Pr[\text{AS}(c_i) > t] \\ &= \prod_{c_i \in W} \left(\sum_{r=t}^n \Pr[\text{AS}(c_i) = r] \right) - \prod_{c_i \in W} \left(\sum_{r=t+1}^n \Pr[\text{AS}(c_i) = r] \right). \end{aligned}$$

Since the SW-DIST problem under the Candidate Probability model can be solved in polynomial time (see Theorem 3.2), computing $\Pr[\text{AS}(c_i) = r]$ can be done in polynomial time, implying that $\text{SWM-PROB}(W, p)$ under the Candidate Probability model is solvable in polynomial time. \square

5 COMMITTEES THAT MAXIMIZE SWM PROBABILITY

We now turn our attention to the problem of $\text{EXISTSSWM-PROB}(p)$, which asks whether there exists a committee W such that W is SWM with probability at least p . Our first result is that under the Lottery model, $\text{EXISTSSWM-PROB}(p)$ is NP-hard, even in the instance with single voter.

THEOREM 5.1. *Under the Lottery model, $\text{EXISTSSWM-PROB}(p)$ is NP-hard, even when $n = 1$.*

PROOF SKETCH. We prove that the problem is NP-hard via a reduction from the MIN- r -UNION (MrU) problem [28]. In MrU, we are given a universe U of m elements, a collection $\mathcal{S} \subseteq 2^U$ of q sets, and two integers $r \leq q$ and ℓ . The goal is to decide whether there exists a sub-collection $I \subseteq [q]$ with size r such that $|\bigcup_{i \in I} S_i| \leq \ell$. In the reduction, our construction maps a YES MrU instance to the existence of a committee W of size ℓ with $\Pr[W \text{ is SWM}] \geq \frac{r}{q}$, and vice versa. \square

Theorem 5.1 implies that $\text{EXISTSSWM-PROB}(p)$ is NP-hard also for the Joint Probability model, as both uncertainty models coincide in single-voter instances. In view of this computational intractability, we consider the special case when $p = 1$, i.e., the $\text{EXISTSSWM-PROB}(1)$ problem, which involves determining whether there exists a committee W that is SWM with probability 1. For ease of clarity, we say a committee W^* is *necessarily* SWM if W^* is SWM with probability 1.

We present Algorithm 3 to show that EXISTS-SWM-PROB(1) is solvable in polynomial time. The algorithm is built upon the “dominance graph”. Specifically, for any candidate pair $(c_i, c_j) \in C$, we say that c_i “dominates” c_j if for every plausible approval profile \mathcal{A} , $AS(c_i, \mathcal{A}) \geq AS(c_j, \mathcal{A})$, denoted by $c_i \succeq^{AS} c_j$. The dominance relation between any candidate pair (c_i, c_j) can be verified in polynomial time by constructing the *deterministic* approval profile $\bar{\mathcal{A}}$ established in Theorem 4.4.

By enumerating all candidate pairs, we construct a dominance digraph $G = (C, E)$ where C is the set of candidates and E contains directed edge representing dominance relations. Specifically, each edge $(c_i, c_j) \in E$ indicates $c_i \succeq^{AS} c_j$. In case $AS(c_i, \mathcal{A}) = AS(c_j, \mathcal{A})$, we break ties by lexicographic order, treating c_i as dominating c_j . No edge is added between c_i and c_j if neither dominates the other. Note that the dominance relation is transitive, and thus the digraph G is acyclic. With the dominance graph constructed, we now present an auxiliary lemma that underpins the design of Algorithm 3.

LEMMA 5.2. *Given a dominance graph $G = (C, E)$, for any $c_i, c_j \in C$ with no edge between c_i and c_j , any necessarily SWM committee W^* satisfies either $\{c_i, c_j\} \subseteq W^*$ or $\{c_i, c_j\} \cap W^* = \emptyset$.*

Lemma 5.2 establishes that, for each candidate pair c_i, c_j without dominance relation, if a necessarily SWM committee exists, then c_i and c_j must either both be included or both be excluded. Leveraging this property, we now present Algorithm 3 to solve the EXISTS-SWM-PROB(1) problem.

Algorithm 3 EXISTS-SWM-PROB(1) algorithm for the Lottery model

Input: $G = (C, E), k$.

Output: YES or NO.

- 1: Initialize $W \leftarrow \emptyset$ and $\bar{G} \leftarrow (\bar{C} = C, \bar{E} = E)$;
 - 2: **while** $|W| \leq k$ **do**
 - 3: Select a candidate c^* with zero indegree in \bar{G} (breaking ties arbitrarily);
 - 4: Add c^* into W ;
 - 5: Update \bar{G} by deleting all the edges in $\{(c^*, c'), c' \in C \setminus \{c^*\} : (c^*, c') \in \bar{E}\}$;
 - 6: **end while**
 - 7: Return YES if $\forall c \in W, c' \in C \setminus W : (c, c') \in E$ otherwise NO;
-

Algorithm 3 takes the dominance graph G as input. First, it initializes an empty candidate set W and creates a duplicate of G , denoted as \bar{G} . The algorithm then iteratively selects a candidate c^* with zero indegree, adds it to W , and updates \bar{G} by removing the outgoing edges from c^* in each round (lines 2-6). Since G is acyclic, a node with zero indegree is guaranteed to exist in the first iteration of the while loop. In each subsequent iteration, \bar{G} remains acyclic as it is a subgraph of G , ensuring that there is always a node with zero indegree, which implies that the algorithm terminates.

THEOREM 5.3. *Under the Lottery model, EXISTS-SWM-PROB(1) is solvable in polynomial time.*

For Candidate Probability model, identifying the complexity of EXISTS-SWM-PROB(p) problem becomes substantially more challenging, even under the restrictive 3VA setting. A natural initial hypothesis is that the expected SWM committee also maximizes the

probability of being SWM. Unfortunately, this intuition does not hold in general. We provide a counterexample involving 3 voters and 4 candidates to illustrate the inherent difficulty.

	1	2	3	4
1	0.5	1.0	1.0	1.0
2	0.5	0.5	1.0	0.5
3	0.5	0.0	0.0	0.0

Each element $p_{i,c}$ is the probability that voter i approves candidate c . We first compute that committees $W_1 = \{1, 3\}$, $W_2 = \{2, 3\}$ and $W_3 = \{3, 4\}$ all achieve the highest expected social welfare, with a value of 2.5. However, their probability of being SWM differ: $\Pr[W_1 \text{ is SWM}] = \frac{19}{32}$, $\Pr[W_2 \text{ is SWM}] = \frac{18}{32}$, and $\Pr[W_3 \text{ is SWM}] = \frac{18}{32}$. This example highlights the inherent complexity and subtlety of the EXISTS-SWM-PROB(p) problem, even under the 3VA model. Despite these challenges, we prove that it is polynomial-time solvable when $n = 1$ or k is constant (See details in the appendix). Beyond these aforementioned results under restrictive assumptions, we also establish a positive result for EXISTS-SWM-PROB(1) problem.

THEOREM 5.4. *Under the Candidate Probability model, EXISTS-SWM-PROB(1) is solvable in polynomial time.*

The core idea is a poly-time algorithm which constructs a profile consisting solely of certain approval ballots and computes the SWM committee for this deterministic approval profile². The algorithm returns YES if the committee is a certificate for YES instance for SWM-PROB($W, 1$), otherwise returns NO. It is worth mentioning that for both Lottery and Candidate Probability model, not only can we determine the EXISTS-SWM-PROB(1) problem in polynomial time, but we can also compute the committee maximizing the probability of being SWM by our proposed polynomial-time algorithms.

6 ROBUST COMMITTEES

In many settings, not only deciding the EXISTS-SWM-PROB problem is intractable, but it can also be highly sensitive to the realizations of the approval profiles. In particular, it might be the case that a committee that maximizes probability of being SWM may only be social welfare maximizing for only a small fraction of the plausible profiles while performing poorly in the remaining plausible profiles. As a result, this motivates us to study *robust* committees. Intuitively, a committee is considered robust if it achieves approximately optimal social welfare with high probability.

Definition 6.1 ((α, β) -Robust Committee). Given any uncertain preference model, denote by W^* the *random committee* that maximizes social welfare under any plausible approval profile and a **random variable** Z representing its social welfare. A committee W is (α, β) -robust if it satisfies

$$\Pr[\text{SW}(W) \geq \alpha \cdot Z] \geq \beta, \text{ where } \alpha, \beta \in (0, 1].$$

Although we know that the a committee maximizing the expected social welfare may not be the committee with highest probability of being SWM, the following result shows that it is guaranteed to be $(\frac{1}{2}, \frac{1}{2})$ -robust.

²Tie-breaking by selecting the committee with the highest number of approvals with positive probabilities.

THEOREM 6.2. *Under the 3VA model, any committee W maximizing the expected social welfare is $(\frac{1}{2}, \frac{1}{2})$ -robust.*

PROOF. Let W be a committee that maximizes the expected social welfare. To prove W is $(\frac{1}{2}, \frac{1}{2})$ -robust, we aim to show that $\Pr[\text{SW}(W) \geq \frac{1}{2} \cdot Z] \geq \frac{1}{2}$, where Z is the random variable representing the social welfare of the social welfare maximizing committee W^* . We next introduce the committee \bar{W} , denoting the committee that has the maximum number of non-zero approval entries (that is, the number of pairs (i, c) for $i \in N$ and $c \in C$ such that $p_{i,c} \neq 0$).

For any plausible approval profile \mathcal{A} under the 3VA model, we first observe that the social welfare achieved by W^* is at most the number of non-zero approval entries for any k candidates in C . Moreover, recall the definition of \bar{W} and we know that the number of non-zero approval entries for any k candidates in C is upper-bounded by $2 \cdot \mathbb{E}[\text{SW}(\bar{W})]$ since \bar{W} has the highest number of non-zero approval entries and $\mathbb{E}[\text{SW}(\bar{W})]$ is equal to the sum value of all approval entries for \bar{W} and the equality holds for the upper bound when all the non-zero approval entries for committee \bar{W} are $\frac{1}{2}$. That is, $\forall c \in \bar{W}, \forall i \in N$ such that $p_{i,c} \neq 0$, then $p_{i,c} = \frac{1}{2}$. Consequently, we have $\text{SW}(W^*) \leq 2 \cdot \mathbb{E}[\text{SW}(\bar{W})]$ for any plausible approval profile \mathcal{A} . On the other hand, by the definition of W , we have $\mathbb{E}[\text{SW}(W)] \geq \mathbb{E}[\text{SW}(\bar{W})]$. Hence, we bound the probability of $\text{SW}(W) \geq \frac{1}{2} \cdot Z$ by

$$\begin{aligned} \Pr[\text{SW}(W) \geq \frac{1}{2} \cdot Z] &= \Pr[\text{SW}(W) \geq \frac{1}{2} \cdot \text{SW}(W^*)] \\ &\geq \Pr[\text{SW}(W) \geq \mathbb{E}[\text{SW}(\bar{W})]] \\ &\geq \Pr[\text{SW}(W) \geq \mathbb{E}[\text{SW}(W)]] = \frac{1}{2}. \end{aligned}$$

The last step holds as $\text{SW}(W)$ follows a shifted binomial distribution, i.e., $\text{SW}(W) \sim \text{Bin}(y, \frac{1}{2}) + x$, thus the probability that $\text{SW}(W)$ is larger than its expectation is $\frac{1}{2}$. Therefore, for any expected social welfare maximization committee W under the 3VA model, we have $\Pr[\text{SW}(W) \geq \frac{1}{2} \cdot Z] \geq \frac{1}{2}$, i.e., W is $(\frac{1}{2}, \frac{1}{2})$ -robust. \square

Since a committee maximizing the expected social welfare can be computed in polynomial time, Theorem 6.2 shows that $(\frac{1}{2}, \frac{1}{2})$ -robust committees can be efficiently computed under the 3VA model. However, our next result shows that a similar result does not hold in the more general setting of the Candidate Probability model.

THEOREM 6.3. *Under the Candidate Probability model, for $\alpha, \beta \in (0, 1]$, no committee is (α, β) -robust.*

PROOF. Given any $\alpha, \beta \in (0, 1]$, we construct an instance under which no committee is (α, β) -robust. Consider an instance with single voter. The sole voter independently approves each candidate with identical probability $p = \frac{\beta}{2}$. Consider the candidate number m such that $m > \frac{\log(p)}{\log(1-p)}$ and set the committee size to be $k = 1$.

Since all candidates have the same approval probability and the committee size is 1, any committee will have the same performance in terms of the probability of being SWM. Without loss of generality, let $W = \{c_1\}$. Recall the definition of random variable Z which represents the social welfare of the SWM committee. In the constructed instance, we express Z as $Z = \max_{c_i \in C} \{\text{AS}(c_i)\}$. For

committee W , we have

$$\Pr[\text{SW}(W) \geq \alpha \cdot Z] = \Pr\left[\text{AS}(c_1) \geq \alpha \cdot \max_{c_i \in C \setminus \{c_1\}} \{\text{AS}(c_i)\}\right]. \quad (3)$$

Now consider $\Pr[\max_{c_i \in C \setminus \{c_1\}} \{\text{AS}(c_i)\} = 1]$, we obtain

$$\begin{aligned} \Pr\left[\max_{c_i \in C \setminus \{c_1\}} \{\text{AS}(c_i)\} = 1\right] &= 1 - \Pr[\forall c_i \in C \setminus \{c_1\}, \text{AS}(c_i) = 0] \\ &= 1 - (1-p)^{m-1}. \end{aligned} \quad (4)$$

Let Y be $\max_{c_i \in C \setminus \{c_1\}} \{\text{AS}(c_i)\}$. Conditioning on Y , we have

$$\begin{aligned} \Pr[\text{SW}(W) \geq \alpha \cdot Z] &= \Pr[\text{AS}(c_1) \geq \alpha \cdot Y] \quad (\text{Equation (3)}) \\ &= \left(\Pr[\text{AS}(c_1) \geq 0 \mid Y = 0] \cdot \Pr[Y = 0]\right) \\ &\quad + \left(\Pr[\text{AS}(c_1) \geq \alpha \mid Y = 1] \cdot \Pr[Y = 1]\right) \\ &= \Pr[Y = 0] + \Pr[\text{AS}(c_1) \geq \alpha] \cdot \Pr[Y = 1] \\ &\quad (\text{Independence}) \\ &= (1-p)^{m-1} + \Pr[\text{AS}(c_1) \geq \alpha] \cdot (1 - (1-p)^{m-1}) \\ &\quad (\text{Equation (4)}) \\ &= (1-p)^{m-1} + p \cdot (1 - (1-p)^{m-1}) \\ &\quad (\text{With probability } p \text{ approving } c_1) \\ &= p + (1-p)^m < \beta. \quad (m > \frac{\log(p)}{\log(1-p)}) \end{aligned}$$

This implies that there is no committee satisfying (α, β) -robust in this instance. \square

Since every candidate probability instance admits a unique representation as a lottery model (Proposition 2.1). Hence, the impossibility result applies for the lottery and joint probability model.

7 CONCLUSIONS

This paper initiates the study of social welfare maximization under uncertainty in approval-based committee voting. Given the relevance of such voting rules in applications like recommender systems and block-chain governance, our framework provides a foundation for incorporating uncertainty into these domains. Many of our results include explicit polynomial-time algorithms, offering a general toolkit for addressing a broader class of set selection problems under uncertainty. Our analysis focuses on social welfare defined by the size of the intersection between the selected committee and voters' approval sets. It would be interesting to investigate whether some of our techniques extend to more general satisfaction functions. A particularly compelling open question is the computational complexity of $\text{EXISTS}_{\text{SWM-PROB}}(p)$ under the candidate probability model, even in the restricted 3-Valued Approval (3VA) setting. Since $\text{EXISTS}_{\text{SWM-PROB}}(p)$ is NP-hard in other models, a promising direction for future work is the development of approximation algorithms for this problem.

ACKNOWLEDGMENTS

This work was partially supported by the NSF-CSIRO grant on "Fair Sequential Collective Decision-Making" (Grant No. RG230833) and the ARC Laureate Project FL200100204 on "Trustworthy AI." Venkateswara Rao Kagita was funded by the Science and Engineering Research Board (SERB) through the SIRE Fellowship (Project ID SIR/2022/001217).

REFERENCES

- [1] Haris Aziz, Péter Biró, Ronald de Haan, and Baharak Rastegari. 2019. Pareto optimal allocation under uncertain preferences: uncertainty models, algorithms, and complexity. *Artificial Intelligence* 276 (2019), 57–78.
- [2] Haris Aziz, Péter Biró, Serge Gaspers, Ronald de Haan, Nicholas Mattei, and Baharak Rastegari. 2020. Stable matching with uncertain linear preferences. *Algorithmica* 82 (2020), 1410–1433.
- [3] Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. 2017. Justified representation in approval-based committee voting. *Social Choice and Welfare* 48, 2 (2017), 461–485.
- [4] H. Aziz, E. Elkind, S. Huang, M. Lackner, L. S. Fernández, and P. Skowron. 2018. On the Complexity of Extended and Proportional Justified Representation. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 902–909.
- [5] Haris Aziz, Isaiah Iliffe, Bo Li, Angus Ritossa, Ankan Sun, and Mashbat Suzuki. 2024. Envy-Free House Allocation under Uncertain Preferences. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*. 9477–9484.
- [6] Haris Aziz, Venkateswara Rao Kagita, Baharak Rastegari, and Mashbat Suzuki. 2024. Approval-Based Committee Voting under Uncertainty. In *Proceedings of the 17th International Symposium of Algorithmic Game Theory (SAGT)*.
- [7] Evripidis Bampis, Konstantinos Dogeas, Thomas Erlebach, Nicole Megow, Jens Schloter, and Amitabh Trehan. 2024. Competitive Query Minimization for Stable Matching with One-Sided Uncertainty. In *Proceedings of the Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 317)*. 17:1–17:21.
- [8] N. Barrot, L. Gourvès, J. Lang, J. Monnot, and B. Ries. 2013. Possible Winners in Approval Voting. In *Proceedings of 3rd International Conference of Algorithmic Decision Theory (ADT)*. 57–70.
- [9] Niclas Boehmer, Markus Brill, Alfonso Cevallos, Jonas Gehrlein, Luis Sánchez-Fernández, and Ulrike Schmidt-Kraepelin. 2024. Approval-Based Committee Voting in Practice: A Case Study of (over-)Representation in the Polkadot Blockchain. *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, 9519–9527.
- [10] Craig Boutilier and Jeffrey S Rosenschein. 2016. Incomplete Information and Communication in Voting. *Handbook of computational social choice* 10 (2016), 223–257.
- [11] Markus Brill, Hayrullah Dindar, Jonas Israel, Jérôme Lang, Jannik Peters, and Ulrike Schmidt-Kraepelin. 2023. Multiwinner voting with possibly unavailable candidates. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*. 5532–5539.
- [12] Markus Brill, Rupert Freeman, Svante Janson, and Martin Lackner. 2024. Phragmén’s voting methods and justified representation. *Mathematical Programming* 203, 1-2 (2024), 47–76.
- [13] Jeff Burdges, Alfonso Cevallos, Peter Czaban, Rob Habermeier, Syed Hosseini, Fabio Lama, Handan Kilinc Alper, Ximin Luo, Fatemeh Shirazi, Alistair Stewart, et al. 2020. Overview of polkadot and its design considerations. *arXiv preprint arXiv:2005.13456* (2020).
- [14] Abhijnan Chakraborty, Gourab K Patro, Niloy Ganguly, Krishna P Gummedi, and Patrick Loiseau. 2019. Equality of voice: Towards fair representation in crowd-sourced top-k recommendations. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT)*. 129–138.
- [15] Virginie Do, Matthieu Hervouin, Jérôme Lang, and Piotr Skowron. 2022. Online Approval Committee Elections. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*. 251–257.
- [16] M. R. Garey and D. S. Johnson. 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness (Series of Books in the Mathematical Sciences)*. W. H. Freeman.
- [17] Grzegorz Gawron and Piotr Faliszewski. 2024. Using multiwinner voting to search for movies. *Theory and Decision* (2024), 1–32.
- [18] D. Halpern, G. Kehne, A. D. Procaccia, J. Tucker-Foltz, and M. Wüthrich. 2023. Representation with incomplete votes. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*. 5657 – 5664.
- [19] N. Hazon, Y. Aumann, S. Kraus, and M. Wooldridge. 2012. On the evaluation of election outcomes under uncertainty. *Artificial Intelligence* 189 (2012), 1–18.
- [20] Aviram Imber, Jonas Israel, Markus Brill, and Benny Kimelfeld. 2025. Approval-based committee voting under incomplete information. *Artificial Intelligence* (2025), 104381.
- [21] Jonas Israel and Markus Brill. 2024. Dynamic proportional rankings. *Social Choice and Welfare* (2024), 1–41.
- [22] D Marc Kilgour. 2010. Approval balloting for multi-winner elections. In *Handbook on approval voting*. Springer, 105–124.
- [23] D. M. Kilgour and E. Marshall. 2012. Approval balloting for fixed-size committees. In *Electoral Systems*. Springer, Chapter 12, 305–326.
- [24] K. Konczak and J. Lang. 2005. Voting procedures with incomplete preferences. In *Multidisciplinary Workshop on Advances in Preference Handling*. 124–129.
- [25] Martin Lackner and Piotr Skowron. 2023. *Multi-Winner Voting with Approval Preferences*. Springer.
- [26] Piotr Skowron, Martin Lackner, Markus Brill, Dominik Peters, and Edith Elkind. 2017. Proportional rankings. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*. 409–415.
- [27] Zoi Terzopoulou, Alexander Karpov, and Svetlana Obraztsova. 2021. Restricted domains of dichotomous preferences with possibly incomplete information. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)*. 5726–5733.
- [28] Staal A Vinterbo. 2002. *A note on the hardness of the k-ambiguity problem*. Technical Report.