A Fairness-Aware Minimum Cost Consensus Model

1st Bapi Dutta
Dept. of Computer Science
Universidad de Jaén
Jaén, Spain
bdutta@ujaen.es

2nd Diego García-Zamora
Dept. of Mathematics
Universidad de Jaén
Jaén, Spain
dgzamora@ujaen.es

3rdRosa M. Rodríguez
Dept. of Computer Science
Universidad de Jaén
Jaén, Spain
rmrodrig@ujaen.es

4th Luis Martínez
Dept. of Computer Science
Universidad de Jaén
Jaén, Spain
martin@ujaen.es

***Abstract*—Minimum cost consensus (MCC) models are widely used in group decision-making to achieve agreement with minimal opinion adjustment costs. Although recent studies have incorporated fairness into MCC frameworks, most existing approaches rely on aggregate social welfare or inequality-based axioms and remain insensitive to subgroup structures induced by protected attributes. As a result, consensus outcomes may still systematically favor certain subgroups despite being cost-optimal or globally fair. To address this limitation, this paper proposes a fairness-aware MCC framework that explicitly accounts for subgroup-level fairness. Fairness is defined through deviation-based measures that evaluate the parity between the final consensus outcome and the opinions of subgroups formed by protected attributes. Fairness is incorporated into MCC models via constraint-based formulations and a learning-inspired objective that interpolates between utilitarian and Rawlsian principles. The proposed models preserve convexity and computational tractability, while enabling a flexible balance between efficiency, consensus, and subgroup fairness in heterogeneous group decision-making environments.**

***Index Terms*—group decision-making, minimum cost consensus model, fairness**

I. Introduction

Group decision-making (GDM) processes are fundamental to organizational management, policy formulation, and collaborative problem-solving across diverse domains [1]. In these processes, multiple decision-makers (DMs) with potentially conflicting preferences must reach a collective agreement that reflects the group's collective wisdom while respecting individual perspectives. Consensus-reaching processes (CRPs) have emerged as a critical mechanism for facilitating such collective decisions, suggesting that DMs to iteratively adjust their opinions until an acceptable level of agreement is achieved [2].

Traditional consensus models have primarily emphasized achieving agreement through opinion convergence, typically assessed using consensus levels or similarity-based metrics. While effective in measuring the extent of agreement, these approaches often overlook the effort required for individuals to adjust their opinions. To address this limitation, minimum cost consensus (MCC) models represent a significant advancement in CRP theory by explicitly considering the costs associated with opinion modifications [3], [4]. These models recognize that opinion adjustments impose psychological, temporal, and resource costs on DMs, and seek to minimize these costs while achieving consensus. The optimization perspective introduced by MCC models has proven particularly valuable in large-scale group decision-making contexts, where the coordination of numerous participants requires efficient resource allocation [5].

Despite the considerable advances in MCC models, it has been found that they often overlook fairness considerations [6], [7]. A consensus solution that minimizes total adjustment costs may disproportionately burden certain DMs, that lead inequitable distributions of modification effort or utility. Therefore, the challenge of integrating fairness considerations into cost-optimized consensus processes has thus emerged as a critical research frontier in group decision-making theory.

The interest in incorporating fairness-related criteria into MCC-based GDM is steadily growing due to the reliance on intelligent group decision systems in several societal aspects [8]–[11]. However, it is not far from obvious how to formulate such ethical concerns in multi-stakeholder scenarios like GDM, because there are multiple views on the notion of fairness [7], [10], [12]. Most of the studies that incorporate fairness in MCC focus on utilitarian ethics in conjunction with social welfare axioms of different economic measures of fairness based on the following human morals (axioms/perceptions/notions): (i) fairness has often been considered as the negation of inequality measure. It is believed that greater fairness can be attained by diminishing inequality in utility distribution [13]; (ii) fairness can often be observed as human self-centered inequity aversion behavior. People experience inequity if they are worse off in material terms than others, and they also feel inequity if they are better off [10]; (iii) fairness can be achieved through a fair negotiation (bargaining) process. It can be achieved by proportionally distributing utility among stakeholders by different bargaining mechanisms in game theory [7], [9], [12], [14].

One common issue with these existing approaches is that such fairness axioms blindfold the demographic and other attributes of the stakeholders and do not warrant equitable shares under such attributes, particularly when decision-making groups contain subgroups defined by protected attributes or distinct characteristics [6]. This oversight can lead to consensus outcomes that systematically favor certain subgroups while disadvantaging others, undermining the legitimacy and acceptability of collective decisions.

To address these challenges, this contribution proposes a unified fairness-aware framework for MCC-based GDM.

Instead of relying solely on abstract social welfare axioms or purely utilitarian cost minimization, we explicitly incorporate subgroup-level fairness concerns induced by protected attributes into the CRP. Fairness is formalized in terms of deviation parity between the final consensus outcome and the opinions of subgroups defined by protected attributes, ensuring that no subgroup is systematically disadvantaged in the consensus result. Within this framework, we develop two complementary fairness integration strategies: (i) a constraint-based approach, where fairness requirements are embedded directly into the MCC model as explicit constraints, and (ii) a learning-inspired objective reformulation that interpolates between utilitarian and Rawlsian principles through a smooth, convex objective function. This dual perspective allows DMs to flexibly balance efficiency, consensus, and fairness, while preserving the tractability and interpretability of MCC models. By grounding fairness in measurable deviation-based criteria and embedding it into optimization-based consensus mechanisms, the proposed approach provides a principled and operational solution to fairness-aware consensus formation in heterogeneous GDM settings.

The remainder of this contribution is organized as follows. Section II introduces the preliminary concepts and notations required for consensus-based GDM and MCC models. Section III formalizes the notion of fairness in GDM, introduces subgroup-based fairness measures associated with protected attributes, and discusses alternative aggregation-based fairness formulations. Section IV integrates these fairness notions into MCC models, presenting both constraint-based fairness-aware MCC formulations and learning-based fair objective functions, along with their theoretical properties. Finally, Section V concludes the contribution and outlines potential directions for future research on fairness-aware consensus mechanisms in large-scale and heterogeneous decision-making environments.

II. Preliminary

Ben-Arieh and Easton [3] were among the first to formalize the MCC framework for CRPs using a mathematical programming perspective. Their approach interprets consensus formation as an optimization task, where the goal is to adjust the experts' original opinions with the smallest possible modification cost. This is achieved by minimizing an objective function that penalizes deviations between the original and revised opinions. At the same time, feasibility constraints are imposed to ensure that each adjusted opinion remains sufficiently close to the group's collective assessment.

$$\begin{aligned} \min \ & \sum_{k=1}^{m} c_k |\hat{o}_k - o_k| \\ \text{s.t.} \ & \Big\{ |\hat{o}_k - \hat{o}| \leq \varepsilon, ; k = 1, 2, ..., m \end{aligned} \tag{MCC}$$

where $(\hat{o}_1, \hat{o}_2, ..., \hat{o}_m)$ denote the modified opinions of the experts, $\hat{o}$ represents the group opinion which can be computed aggregating individual opinions through aggregation function, and ε specifies the maximum allowed deviation from the collective opinion.

Although MCC-based approaches ensure that the adjusted opinions are close to the collective outcome while minimizing adjustment costs, they do not explicitly incorporate classical consensus indicators commonly used in CRP studies [15]. As a consequence, these models do not explicitly guarantee that a predefined agreement level is reached. To overcome this limitation, Labella et al. [16] proposed the Consensus-based Minimum Cost Consensus (CMCC) model, which integrates an explicit consensus requirement into the optimization framework:

$$\begin{aligned} \min \ & \sum_{k=1}^{m} c_k |\hat{o}_k - o_k| \\ \text{s.t.} \ & \begin{cases} \hat{o} = \sum_{k=1}^{m} w_k \hat{o}_k \\ |\hat{o}_k - \hat{o}| \leq \varepsilon, ; k = 1, 2, ..., m \\ \kappa(\hat{o}_1, \hat{o}_2, ..., \hat{o}_m) \geq \mu_0 \end{cases} \end{aligned} \tag{CMCC}$$

where $\kappa : [0,1]^m \to [0,1]$ is a function that quantifies the overall consensus degree among experts, and $\mu_0 \in [0,1]$ denotes the minimum required consensus level. Here, the group opinion $\hat{o}$ is computed from the individual opinions via weighted arithmetic mean.

III. Fairness in GDM

A. Concepts and formalization

Let $\mathcal{O} = (o_1, o_2, ..., o_m) \in [a,b]^m$ be the opinions of the m DMs $\mathcal{G} = \{dm_1, \ldots, dm_m\}$. Let $\mathcal{F} = \{F_1, F_2, \ldots F_p\}$ be the *protected attributes* of the DMs associated with the fairness concerns. By *protected attribute*, we essentially mean an attribute that can divide the group $\mathcal{G}$ into sub-groups [17]. For instance, when the group is formed to analyze the diverse perspectives regarding decision problems, the protected attribute could be DMs' background or demographic features. By fairness concerns, we essentially mean that the group decision under a consensus agreement should not favor any subgroup dictated by the protected attribute. Usually, the consensus condition provides some sort of agreement among the involved DMs in the group. However, such an overall agreement in the group does not necessarily imply the closeness of each subgroup decision formed by the protected attribute with the group decision, and eventually, it may favor a subgroup with a specific protected attribute.

There might be many ways to define the notion of *fairness* based on the decision scenario at hand. Here, we introduce the notion of fairness in terms of deviation, i.e., how the final decision outcome deviates from the decision of the individuals of a subgroup indicated by the protected attribute. If all subgroups $\mathcal{G}_{F_s} = \{dm_k : dm_k$ has protected attribute $F_s\}$ deviate equally from the final decision outcome, we will refer to the outcome of the decision-making process as fair.

B. Fairness measure:

To operationalize these fairness concerns within a consensus-based GDM framework, it is necessary to quantify how fairly the final outcome reflects the opinions of different subgroups. We therefore introduce a subgroup-level fairness measure that captures the deviation between the final decision outcome and the modified opinions of DMs sharing the same protected attribute. In this context, we assume that the DMs are willing to modify their initial opinions in the subsequent consensus process to reach a fair solution. Let $\hat{\mathcal{O}} = (\hat{o}_1, \hat{o}_2, ..., \hat{o}_m) \in [a, b]^m$ be the modified opinions of the experts of $\mathcal{G}$ with the final decision outcome $\hat{o}$, and $\mathbb{I}_{\mathcal{G}_{F_s}}$ be the index set of the DMs in the subgroup $\mathcal{G}_{F_s}$. Now onward, the modified opinion vectors of the subgroups $\mathcal{G}_{F_s}$ will be denoted by $\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}$. Let $\mathcal{D} : [0, 1]^2 \to [0, 1]$ be a distance measure over the decision space of DMs that measures the deviation between any two opinions. Keeping in mind these notations, we define the *fairness* measure of the sub-group corresponding to the protected attribute F_s as the mapping $\mathcal{M} : [0, 1]^{|\mathcal{G}_{F_s}|+1} \to [0, 1]$ that measure the distance between final decision outcome and individual decision of the subgroups. When the value of the fairness measure $\mathcal{M}$ approaches zero, the final decision outcome is deemed to be fair to the subgroup $\mathcal{G}_{F_s}$. On the other hand, the values close to 1 indicate that the final decision outcome is completely unfair to the group. Mathematically, this measure should follow certain properties:

- if $\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}} = (\tilde{o}, \tilde{o}, \ldots, \tilde{o})$ then $\mathcal{M}(\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}, \hat{o}) = 0$
- if $\hat{\mathcal{O}}_{\sigma(\mathbb{I}_{\mathcal{G}_{F_s}})}$ is a permutation of the $\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}$, then $\mathcal{M}(\hat{\mathcal{O}}_{\sigma(\mathbb{I}_{\mathcal{G}_{F_s}})}, \hat{o}) = \mathcal{M}(\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}, \hat{o})$

The above-mentioned properties ensure that the fairness measure assigns zero unfairness to a unanimously aligned subgroup and remains invariant to any permutation of subgroup members, thereby capturing both consistency and anonymity in the decision-making process.

One typical way to construct such a measure $\mathcal{M}$ is to use an aggregation function $\mathcal{A} : [a, b]^{|\mathcal{G}_{F_s}|} \to [a, b]$ coupled with distance and can be defined as follows:

$$\mathcal{M}(\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}, \hat{o}) = \mathcal{A}(\mathcal{D}(\hat{o}_{i_1}, \hat{o}), \ldots, \mathcal{D}(\hat{o}_{i_t}, \hat{o})), i_1, \ldots i_t \in \mathbb{I}_{\mathcal{G}_{F_s}} \tag{1}$$

If we consider the collective fairness of the subgroup $\mathcal{G}_{F_s}$, one possible choice could be the *arithmetic mean* operator and in that case Eq. (1) reduced to

$$\mathcal{M}(\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}, \hat{o}) = \frac{1}{|\mathbb{I}_{\mathcal{G}_{F_s}}|} \sum_{i \in \mathbb{I}_{\mathcal{G}_{F_s}}} \mathcal{D}(\hat{o}_{i_1}, \hat{o}) \tag{2}$$

When it is essential that final decision outcomes be fair to each individual of subgroup to avoid a little favoritism to some members of the subgroup (i.e., some are very close to the final outcome while others may be a little further), in such cases, one could choose non-compensatory aggregation function such as $\max$ operator and measure of fairness Eq. (1) becomes

$$\mathcal{M}(\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}, \hat{o}) = \max_{i \in \mathbb{I}_{\mathcal{G}_{F_s}}} \mathcal{D}(\hat{o}_i, \hat{o}) \tag{3}$$

Another way to measure fairness is from the subgroup's final decision point of view, i.e., how the collective decision of the sub-group differs from the final decision outcome of the group. Taking into account Eq. (1), the fairness measure can be defined as follows:

$$\mathcal{M}(\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}, \hat{o}) = \mathcal{D}(\mathcal{A}(\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}), \hat{o}) \tag{4}$$

where $\mathcal{A} : [a, b]^{|\mathcal{G}_{F_s}|} \to [a, b]$ is the averaging aggregation function, used to form the collective sub-group opinion.

IV. Fairness in MCC

While MCC models provide an efficient mechanism for reaching agreement by minimizing opinion adjustment costs, they do not explicitly account for fairness considerations among DMs or subgroups. As discussed in the previous section, consensus at the group level does not necessarily imply equitable treatment across subgroups defined by protected attributes. To address this limitation, this section integrates deviation-based fairness notions into the MCC framework. We present two complementary strategies for bias mitigation: a constraint-based approach that enforces subgroup fairness explicitly, and a learning-inspired objective reformulation that balances efficiency and fairness within the consensus process.

A. Bias mitigation through constraints

To ensure that such a model can take care of the fairness concerns of DMs, we embed fairness measures associated with subgroups formed by protected attributes, and the refined model is formulated as follows:

$$\min \sum_{i=1}^{n} c_i |\hat{o}_i - o_i| \quad s.t. \begin{cases} \hat{o} = \mathcal{P}(\hat{o}_1, \ldots, \hat{o}_n) \\ |\hat{o}_i - \hat{o}| \leq \varepsilon, i = 1, \ldots, n \\ \sum_{i=1}^{n} |\hat{o}_i - \hat{o}| \leq \gamma \\ \mathcal{M}(\hat{\mathcal{O}}_{\mathbb{I}_{\mathcal{G}_{F_s}}}, \hat{o}) \leq \delta, F_s \in \mathcal{F} \end{cases} \tag{Fair-MCC}$$

where $\mathcal{P} : [a, b]^n \to [a, b]$ is an averaging aggregation function and $\delta \geqslant 0$ is a control parameter for fairness. The value of δ close to 0 enforces a high fairness degree for all subgroups. It is also possible that different subgroups could seek different levels of fairness degree, and such a scenario could be handled by associating $\delta_{F_s} \geqslant 0$ to each subgroup $\mathcal{G}_{F_s}$.

Typically, the above model considers the fairness aspects in the consensus-based decision-making process. However, it is possible to consider only the fairness aspect, i.e., we are only concerned with reaching a fair solution without bothering about the agreement in the group. If all the subgroups find that the solution is fair for them, we can assume that they have accepted the final decision, and that indicates some level of acceptance agreement. Based on our notion of fairness measure, the idea of finding a fair solution can be translated

into a final decision that differs almost equally from all subgroups' decisions. It can be formally defined as follows:

$$\min \delta \quad s.t. \begin{cases} \hat{o} = \mathcal{P}(o_1, \ldots, o_n) \\ \mathcal{M}(\mathcal{O}_{\mathbb{I}_{\mathcal{G}_{F_s}}}, \hat{o}) \leq \delta, F_s \in \mathcal{F} \\ \delta \geq 0 \end{cases} \tag{Fair-M}$$

The model works as an implementation "min-max" principle, where we attempt to minimize the worst fairness of the subgroup. Note that here we do not consider the modification of the opinions and thus the notion of concepts. It depends heavily on the initial opinions of DMs. The value of δ could be high when subgroups' opinions are very diverse. There is no way to ensure that we can reach the desired fairness level. However, it is the best, fair solution for each subgroup given the initial opinion.

Note that we could extend this notion of fairness to the traditional feedback mechanism-based CRP by guiding the opinion modification process grounded on the feedback and fairness measure.

The notion of the fairness measure introduced in Eq. (1) could be viewed as a cardinal measure. It can be easily extended to deal with multiple alternative assessment problems in terms of pairwise comparisons or multi-criteria problems. Moreover, this measure could further be extended to the ordinal settings, where we define fairness by comparing the parity of the ranking orders between the subgroups.

B. Bias mitigation through learning principle

Let us consider the typical MCC model with the equal cost of changing the opinions as follows:

$$\min \sum_{i=1}^{n} |\hat{o}_i - o_i| \quad s.t. \begin{cases} \hat{o} = \mathcal{P}(\hat{o}_1, \ldots, \hat{o}_n) \\ |\hat{o}_i - \hat{o}| \leq \varepsilon, i = 1, \ldots, n \end{cases} \tag{U-MCC}$$

From the utilitarian perspective [18], the model seeks to minimize the sum of total changes in the DMs' initial opinions. Assuming that the utility of DMs is proportional to their changes, one could find that the MCC model attempts to maximize the group utility. The fairness of such allocation comes from the overall group benefits. However, the mechanism could hurt one or more individuals for the greater benefit of others in the group. That could bring inequality for members of the group and question the fairness protocol employed in the model.

Such indifference to some members of the group could be solved through Rawlsian "Difference principle", which is more concerned with individual inequalities [19]. The principle argued that the inequalities should be distributed in a manner such that it brings the greatest benefit to the least advantaged. It could be modeled through the min-max principle as follows:

$$\min \max_i |\hat{o}_i - o_i| \quad s.t. \begin{cases} \hat{o} = \mathcal{P}(\hat{o}_1, \ldots, \hat{o}_n) \\ |\hat{o}_i - \hat{o}| \leq \varepsilon, i = 1, \ldots, n \end{cases} \tag{R-MCC}$$

While the model (R-MCC) is more concerned about individual fairness, it is not concerned about the majority of the DMs.

Both approaches have some shortcomings. Individual fairness can be considered an endpoint within the spectrum of group fairness. However, it tends to have limitations in terms of generalizability and can be challenging to quantify accurately. Conversely, group fairness may overlook disparities in outcomes within a given group. To this end, we attempt to build an objective function that seemingly interpolates these two conflicting paradigms of individual and group fairness. The MCC model with a fair objective could be represented as follows:

$$\min \ \frac{1}{\theta} \log \big(\sum_{i=1}^{n} e^{\theta |\hat{o}_i - o_i|} \big) \quad s.t. \begin{cases} \hat{o} = \mathcal{P}(\hat{o}_1, \ldots, \hat{o}_n) \\ |\hat{o}_i - \hat{o}| \leq \varepsilon, i = 1, \ldots, n \end{cases} \tag{I-MCC}$$

As $\theta \to \infty$, the sum in the objective of (I-MCC) is dominated by the term with maximum deviation from the initial opinions of the DM and, thus, approaches the Rawlsian minimax objective of the model (R-MCC). On the other hand, when $\theta \to 0$, the exponential is approximately linear in its arguments, and eventually, it leads to the utilitarian objective of the model (U-MCC). Note that the model (I-MCC) treats the group as homogeneous. The parameter θ could be interpreted as fairness concerns. The value close to 0 represents more DMs' concern about overall group benefits. Conversely, the high value of θ indicates DM is more concerned about individual fairness.

Theorem 1. The objective function of the model (I-MCC) is convex for all $\theta \in (0, \infty)$.

Proof. Let $\hat{\mathcal{O}}^1 = (\hat{o}_1^1, \ldots, \hat{o}_n^1) \in \mathcal{X}$ and $\hat{\mathcal{O}}^2 = (\hat{o}_1^2, \ldots, \hat{o}_n^2) \in \mathcal{X}$ be any two point from the feasible region of optimization model (I-MCC), $\mathcal{X}$. Further, we denote the objective function as $f(\hat{\mathcal{O}}, \theta)$. For any $\lambda_1, \lambda_1 \in [0, 1]$ and $\lambda_1 + \lambda_2 = 1$, we have

$$\theta |(\lambda_1 \hat{o}_i^1 + \lambda_2 \hat{o}_i^2) - \hat{o}_i| \leq \theta\lambda |\hat{o}_i^1 - \hat{o}_i| + \theta\lambda_2 |\hat{o}_i^2 - \hat{o}_i| \ \forall \ i = 1, \ldots n.$$

Since $\log(.)$ is an increasing function, we obtain

$$\log \big(\sum_{i=1}^{n} e^{\theta |(\lambda_1 \hat{o}_i^1 + \lambda_2 \hat{o}_i^2) - o_i|} \big) \leq \log \big(\sum_{i=1}^{n} e^{\theta \lambda_1 |\hat{o}_i^1 - o_i|} \ e^{\theta \lambda_2 |\hat{o}_i^2 - o_i|} \big).$$

Taking $a_i = e^{\theta |\hat{o}_i^1 - o_i|}$ and $b_i = e^{\theta |\hat{o}_i^2 - o_i|}$ and applying Hölder's inequality, we obtain

$$\log \big(\sum_{i=1}^{n} a_i^{\lambda_1} \ b_i^{\lambda_2} \big) \leq \log \Big(\big(\sum_{i=1}^{n} a_i^{\lambda_1 . 1/\lambda_1} \big)^{\lambda_1} \big(\sum_{i=1}^{n} a_i^{\lambda_2 . 1/\lambda_2} \big)^{\lambda_2} \Big) \leq \lambda_1 \log \big(\sum_{i=1}^{n} a_i \big) + \lambda_2 \log \big(\sum_{i=1}^{n} b_i \big)$$

Therefore,

$$\frac{1}{\theta}\log\Big(\sum_{i=1}^{n} e^{\theta\lambda_1|\hat{o}_i^1-o_i|}\, e^{\theta\lambda_2|\hat{o}_i^2-o_i|}\Big) \leq$$
$$\lambda_1\frac{1}{\theta}\log\Big(\sum_{i=1}^{n} e^{\theta|\hat{o}_i^1-o_i|}\Big)+\lambda_2\frac{1}{\theta}\log\Big(\sum_{i=1}^{n} e^{\theta|\hat{o}_i^2-o_i|}\Big)$$

Hence,

$$f(\lambda_1\hat{\mathcal{O}}^1+\lambda_2\hat{\mathcal{O}}^2,\theta)\leq\lambda_1 f(\hat{\mathcal{O}}^1,\theta)+\lambda_2 f(\hat{\mathcal{O}}^2,\theta).$$

□

Corollary 1. Since the objective function of the optimization model (I-MCC) is convex, if the aggregation operator $\mathcal{P}:[a,b]^n\to[a,b]$ to form the group opinion is affine, then the model (I-MCC) is a convex programming problem.

Theorem 2 (Existence of optimal solution for (I-MCC))**.** Let $O=(o_1,\ldots,o_n)\in[a,b]^n$ be the vector of initial opinions. Assume that:

1) the opinion space $[a,b]^n\subset\mathbb{R}^n$ is compact;
2) the aggregation operator $\mathcal{P}:[a,b]^n\to[a,b]$ is continuous;

Then, for any $\varepsilon>0$ and $\theta\in(0,\infty)$, the (I-MCC) model admits at least one optimal solution.

The result follows directly from the Weierstrass theorem, since the feasible set of (I-MCC) is compact, being a closed subset of $[a,b]^n$ due to the continuity of $\mathcal{P}$ and the constraints, and the objective function is continuous for all $\theta\in(0,\infty)$.

Theorem 2 guarantees the well-posedness of the (I-MCC) model independently of convexity assumptions. Convexity becomes relevant only for uniqueness and computational tractability, not for the existence of optimal solutions.

Example 1. In the aim of understanding the impact of the parameter θ, we consider a GDM scenario with five DMs and initial opinions $\mathcal{O}=(2,1.5,3,2,4)$. For the fixed value of the distance between group and individual opinion $\varepsilon=0.3$, we obtained the modified opinions of DMs for different values of θ from (I-MCC) model as well as from the utilitarian and Rawlsian models and depicted Fig. 1. We observe that for the compensation mechanism, the utilitarian model forces the change of DM5 (1.8 units from the initial opinion) more than others to minimize the group changes, whereas the Rawlsian model, being more concerned with individual changes, made the least change for DM5 (0.95 units from the initial opinion). When the values of θ are close to zero, modified opinions of the DMs are almost similar to utilitarian opinions $\hat{\mathcal{O}}=(2,1.85,2.45,2,2.45)$. With the increase of values of θ, the modified opinions are shifted towards the Rawlsian solution $\hat{\mathcal{O}}=(2.45,2.45,3.05,2.75,3.05)$.

When the group is heterogeneous, and each subgroup, against protected feature $F_s\in\mathcal{F}$, is concerned about fairness,

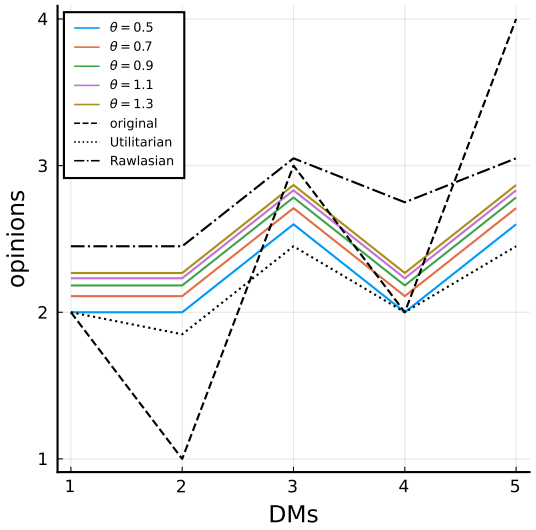

Fig. 1. Opinions modification for different values of θ

the (I-MCC) model can be put into the following form by modifying the fair objective:

$$\begin{aligned}&\min\ \frac{1}{\theta}\log\Big(\sum_{F_s\in\mathcal{F}} e^{\theta\sum_{i\in F_s}|\hat{o}_i-o_i|}\Big)\\ &s.t.\begin{cases}\hat{o}=\mathcal{P}(\hat{o}_1,\ldots,\hat{o}_n)\\ |\hat{o}_i-\hat{o}|\leq\varepsilon, i=1,\ldots,n\end{cases}\end{aligned}\qquad\text{(GI-MCC)}$$

Here, we focus on subgroups' fairness rather than individual fairness in model (I-MCC). When $\theta\to 0$, the objective of (GI-MCC) becomes, $\min\sum_{F_s\in\mathcal{F}}\sum_{i\in F_s}|\hat{o}_i-o_i|$. This indicates that the fairness of the subgroups is taken into account as compensatory ways that hurting some subgroups could be possible for the sake of improving others and eventually overall. Conversely, when $\theta\to\infty$, the objective of (GI-MCC) approaches $\min\max_{F_s\in\mathcal{F}}\sum_{i\in F_s}|\hat{o}_i-o_i|$, which indicates the concern of the least advantaged subgroup's fairness. Thus, the model focuses heavily on subgroups' fairness. Intermediate values of the θ allow us to resolve somewhat both group and subgroup concerns.

The above theoretical results confirm the soundness of the proposed fairness-aware MCC formulations. The existence of optimal solutions ensures that fairness-sensitive objectives are well-defined, while convexity under mild conditions guarantees computational tractability. Together, these properties provide a solid foundation for fair and subgroup-aware consensus formation in heterogeneous group decision-making settings.

V. Conclusion

This paper presented a fairness-aware MCC in GDM, addressing the limitation of traditional MCC models that focus solely on efficiency while neglecting equity among DMs. Fairness was formalized using deviation-based measures that capture how the final consensus outcome differs from the opinions of individuals and subgroups defined by protected attributes.

Two complementary strategies were proposed to incorporate fairness into MCC models. The first embeds subgroup fairness requirements as explicit constraints within the consensus optimization process, ensuring equitable treatment of protected subgroups. The second reformulates the MCC objective through a smooth, convex function that interpolates between utilitarian and Rawlsian fairness principles, enabling a flexible trade-off between overall efficiency and individual or subgroup equity. Theoretical analysis confirmed the convexity of the proposed objective under mild conditions, preserving computational tractability.

The proposed framework is general and adaptable, and can be extended to heterogeneous groups, multiple protected attributes, and more complex decision structures. Future work will focus on dynamic fairness-aware consensus processes, learning fairness parameters from data, and applications in large-scale and AI-supported GDM systems.

ACKNOWLEDGMENTS

This work is supported by the Spanish Ministry of Science, Innovation and Universities, and the Spanish State Research Agency through the Ramón y Cajal Research Grant (RYC2023-045020-I) and Knowledge Generation Projects Grant (PID2024-161073NB-I00), Spain. Generative AI tool has been used only to polish the language during the preparation of the manuscript.

REFERENCES

[1] D. Black, "On the rationale of group decision-making," *Journal of political economy*, vol. 56, no. 1, pp. 23–34, 1948.

[2] F. Herrera, E. Herrera-Viedma *et al.*, "A model of consensus in group decision making under linguistic assessments," *Fuzzy sets and Systems*, vol. 78, no. 1, pp. 73–87, 1996.

[3] D. Ben-Arieh and T. Easton, "Multi-criteria group consensus under linear cost opinion elasticity," *Decision Support Systems*, vol. 43, no. 3, pp. 713–721, 2007.

[4] G. Zhang, Y. Dong, Y. Xu, and H. Li, "Minimum-cost consensus models under aggregation operators," *IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans*, vol. 41, no. 6, pp. 1253–1261, 2011.

[5] Y. Liang, T. Zhang, Y. Tu, and Q. Zhao, "A group consensus reaching model balancing individual satisfaction and group fairness for distributed linguistic preference relations," *Applied Intelligence*, vol. 54, no. 24, pp. 12 697–12 724, 2024.

[6] H. Hosseini, "The fairness fair: Bringing human perception into collective decision-making," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 20, 2024, pp. 22 624–22 631.

[7] F. Meng, D. Zhao, and X. Zhang, "A fair consensus adjustment mechanism for large-scale group decision making in term of gini coefficient," *Engineering Applications of Artificial Intelligence*, vol. 126, p. 106962, 2023.

[8] Z.-S. Chen, Z. Zhu, Z.-J. Wang, and Y. Tsang, "Fairness-aware large-scale collective opinion generation paradigm: A case study of evaluating blockchain adoption barriers in medical supply chain," *Information Sciences*, vol. 635, pp. 257–278, 2023.

[9] F. Jing and X. Chao, "Fairness concern: An equilibrium mechanism for consensus-reaching game in group decision-making," *Information Fusion*, vol. 72, pp. 147–160, 2021.

[10] W.-C. Zou, S.-P. Wan, and S.-M. Chen, "A fairness-concern-based linmap method for heterogeneous multi-criteria group decision making with hesitant fuzzy linguistic truth degrees," *Information Sciences*, vol. 612, pp. 1206–1225, 2022.

[11] B. Dutta and L. Martínez, "Exploring new formulations for distributive fairness in intelligent decision-making," in *Digital Transformation in Artificial Systems*. Elsevier, 2026, pp. 35–52.

[12] Z. Gong, W. Guo, and R. Słowiński, "Transaction and interaction behavior-based consensus model and its application to optimal carbon emission reduction," *Omega*, vol. 104, p. 102491, 2021.

[13] G. Gong, K. Li, and Q. Zha, "A maximum fairness consensus model with limited cost in group decision making," *Computers & Industrial Engineering*, vol. 175, p. 108891, 2023.

[14] F. Meng, J. Tang, and Q. An, "Cooperative game based two-stage consensus adjustment mechanism for large-scale group decision making," *Omega*, vol. 117, p. 102842, 2023.

[15] R. M. Rodríguez, Á. Labella, B. Dutta, and L. Martínez, "Comprehensive minimum cost models for large scale group decision making with consistent fuzzy preference relations," *Knowledge-Based Systems*, vol. 215, p. 106780, 2021.

[16] Á. Labella, H. Liu, R. Rodríguez, and L. Martínez, "A cost consensus metric for consensus reaching processes based on a comprehensive minimum cost model," *European Journal of Operational Research*, vol. 281, pp. 316–331, 2020.

[17] M. Kearns, S. Neel, A. Roth, and Z. S. Wu, "Preventing fairness gerrymandering: Auditing and learning for subgroup fairness," in *International conference on machine learning*. PMLR, 2018, pp. 2564–2572.

[18] J. S. Mill, "Utilitarianism," in *Seven Masterpieces of Philosophy*. Routledge, 2001, pp. 337–383.

[19] J. Rawls, "Atheory of justice," *Cambridge (Mass.)*, 1971.