The *Before, During and After* of Multi-Robot Deadlock

Jaskaran Grover[1], Changliu Liu [1] and Katia Sycara[1]

Abstract

Collision avoidance for multirobot systems is a well studied problem. Recently, control barrier functions (CBFs) have been proposed for synthesizing controllers that guarantee collision avoidance and goal stabilization for multiple robots. However, it has been noted that reactive control synthesis methods (such as CBFs) are prone to *deadlock*, an equilibrium of system dynamics that causes the robots to stall before reaching their goals. In this paper, we analyze the closed-loop dynamics of robots using CBFs, to characterize controller parameters, initial conditions and goal locations that invariably lead the system to deadlock. Using tools from duality theory, we derive geometric properties of robot configurations of an N robot system once it is in deadlock and we justify them using the mechanics interpretation of KKT conditions. Our key deductions are that 1) system deadlock is characterized by a force-equilibrium on robots and 2) deadlock occurs to ensure safety when safety is at the brink of being violated. These deductions allow us to interpret deadlock as a subset of the state space, and we show that this set is non-empty and located on the boundary of the safe set. By exploiting these properties, we analyze the number of admissible robot configurations in deadlock and develop a provably-correct decentralized algorithm for deadlock resolution to safely deliver the robots to their goals. This algorithm is validated in simulations as well as experimentally on Khepera-IV robots.

Keywords

Collision Avoidance, Optimization and Optimal Control, Duality Theory

1 Introduction

Multirobot systems have been studied extensively for solving a variety of complex tasks such as target search Kantor et al. [2003], Grover et al. [2020*b*], sensor coverage Cortes et al. [2004], environmental exploration Burgard et al. [2005] and perimeter guarding Feng and Yu [2019]. Global coordinated behaviors result from executing local control laws on individual robots interacting with their neighbors Ogren et al. [2002], Olfati-Saber et al. [2007]. Typically, the local controllers running on these robots are a combination of a task-based controller responsible for completion of a primary objective and a reactive collision avoidance controller for ensuring safety. However, augmenting the primary task based control with a hand-engineered safety control no longer guarantees that the original task will be satisfied Borrmann et al. [2015]. This problem becomes all the more pronounced when the number of robots increases. Motivated by this drawback, our paper focuses on an algorithmic analysis of the performance-safety trade-offs that result from augmenting a task-based controller with collision avoidance constraints as done using CBF based quadratic programs (QPs) Ames et al. [2017]. Although CBF-QPs mediate between safety and performance in a rigorous way, yet ultimately they are distributed local controllers. Such approaches exhibit a lack of look-ahead, which causes the robots to be trapped in *deadlocks* as noted in Petti and Fraichard [2005]; O'DONNELL [1989]; Wang et al. [2017].

In deadlock, the robots stop while still being away from their goals and persist in this state unless intervened. This occurs because robots reach a state where conflict
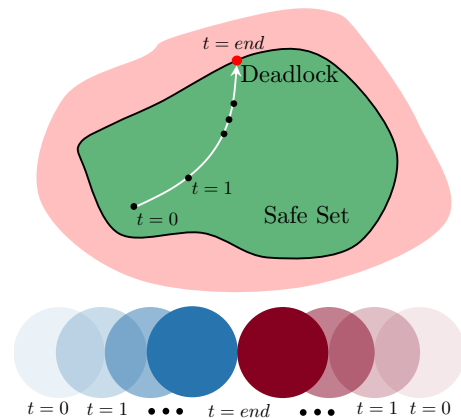
Figure 1. (lower) In the work space, two robots moving towards each other fall in deadlock. (upper) In the state space, the system state converges to the boundary of safe set.

becomes inevitable, *i.e.* a control favoring goal stabilization will violate safety (see the red dot in Fig. 1). Hence, the only feasible strategy is to remain static. Although small perturbations can steer the system away from deadlock, there

is no guarantee that deadlock will not relapse. To circumvent these issues, this work addresses the following technical questions:

1. What initial conditions/controller parameters invariably lead a system to deadlock?
2. What are the salient geometric properties of a system in deadlock?
3. What are all the admissible configurations of robots in deadlock?
4. How can we leverage this information to provably exit deadlock using decentralized controllers?

To address these questions, we first recall technical definitions for CBF-QPs in section 3. We divide our analysis intro three sections consistent with the chronology of deadlock incidence.

1. **Before Deadlock** (section 4) In this section, we analyze the closed-loop dynamics of the multirobot system before the incidence of deadlock to characterize initial conditions and controller parameters that result in deadlock. We show that geometric symmetry in initial conditions and goals locations necessarily results in deadlock and that heterogeneity in robots' controller parameters is not sufficient to avoid deadlock.
2. **During Deadlock** (section 5) In this section, we use duality theory to derive geometric properties of robots' configurations once the system is in deadlock. We propose a novel set-theoretic definition of deadlock to interpret it as a subset of the system's state space with an eye towards devising mitigative controllers to exit this set. We show that system deadlock is characterized by a force equilibrium on robots and that the deadlock set is non-empty and located on the boundary of the system's safe set. Using graph enumeration, we show that the number of geometric configurations of robots that are admissible in deadlock increases combinatorially with the number of robots.
3. **After Deadlock** (section 6) In this section, we design provably-correct decentralized controllers to make the robots exit the deadlock set. We demonstrate this strategy on two and three robots in simulation, and experimentally on Khepera-IV robots.

Finally, we conclude in section 7 with a summary of our work and directions for future research questions.

Remark 1: In the narrative, we follow the technical statement of a theorem/lemma with an intuitive explanation to convey the high-level idea. The proofs are deferred to the appendix to sustain the flow of essential ideas in the main body of the paper.

2   Prior Work

Several existing methods provide inspiration for the results presented here. Among these, two are especially relevant: in the first category, we describe prior methods for collision avoidance and in the second, we focus on deadlock resolution.

Prior Work on Avoidance Control

Avoidance control is a well-studied problem with immediate applications for planning collision-free motions for multi-robot systems. Classical avoidance control assumes a worst case scenario with no cooperation between robots Leitmann and Skowronski [1977, 1983] Cooperative collision avoidance is explored in Stipanović et al. [2007]; Hokayem et al. [2010] where avoidance control laws are computed using value functions. Velocity obstacles have been proposed in Fiorini and Shiller [1998] for motion planning in dynamic environments. They select avoidance maneuvers outside of robot's velocity obstacles to avoid static and moving obstacles by means of a tree-search. While this method is prone to undesirable oscillations, the authors in Van den Berg et al. [2008]; Van Den Berg et al. [2011]; Wilkie et al. [2009] proposed reciprocal velocity obstacles that are immune to such oscillations. Recent work by Douthwaite et al. [2018] gave a comparative analysis of the variants of velocity obstacles and empirically showed incidence of deadlock in symmetric situations. Additionally, control barrier function based controllers have been used in Borrmann et al. [2015]; Wang et al. [2017] to mediate between safety and performance using QPs. Pan et al. [2020] presents a multi-robot feedback control policy augmented with a global planner for robust safe navigation in a complex workspace.

Prior Work on Deadlock Resolution

The importance of coordinating motions of multiple robots while simultaneously ensuring safety, performance and deadlock prevention has been acknowledged in works as early as in O'DONNELL [1989]. Here, authors proposed scheduling algorithms to asynchronously coordinate motions of two manipulators to ensure that their trajectories remain collision free and deadlock free. In the context of mobile robots, Yamaguchi [1999] identified the presence of deadlocks in a cooperative scenario using mobile robot troops. To the best of our knowledge, Jager and Nebel [2001] were the first to propose algorithms for deadlock resolution specifically for multiple mobile robots. Their strategy for collision avoidance modifies planned paths by inserting idle times and resolves deadlocks by asking the trajectory planners of each robot to plan an alternative trajectory until deadlock is resolved. Duhaut et al. [2007] show that without sharing any information on the global situation of the multi-robot system, deadlock cannot be avoided. Further, they proposed a random move + priority-based rule for deadlock resolution. Authors in Li et al. [2005] proposed coordination graphs to resolve deadlocks in robots navigating through narrow corridors. Wang et al. [2017]; Rodríguez-Seda et al. [2016] added perturbation terms to their controllers for avoiding deadlock. Authors in Douthwaite et al. [2019, 2018]; Douthwaite [2019] gave empirical simulation based evidence of incidence of deadlock in symmetric situations and prevented deadlock occurrence by adding small perturbations, whereas our work gives analytical proofs why those geometric arrangements are prone to deadlock with extensions to heterogeneous robots. Zhou et al. [2017a,b] proposed a distributed algorithm to avoid collisions and deadlocks by stopping robots preemptively before these events occur. Alonso-Mora

et al. [2018] presents an algorithm for mission and motion planning for small teams of robots performing a task around moving obstacles guaranteeing safety and providing an algorithm and conditions for deadlock resolution.

In these works, controllers for task-completion and collision avoidance are highly coupled. However, in our work, we assume that a nominal task based control is already given and must be followed as much as possible while maintaining safety using the framework of CBF-QPs. Thus, the mediation between safety and task performance is hierarchical with safety constraints taking precedence over task-satisfaction. Due to this nature, our analysis is generalizable to deadlocks resulting from an arbitrary task-based control and not restricted to just the go-to-goal task. Further, differently from prior work, we characterize analytical properties of the system dynamics to characterize the reasons behind deadlock incidence. We demonstrate that intuitive explanations for geometric properties of a system in deadlock are indeed recovered using duality. Moreover, our analysis can be extended to reveal bottlenecks of any optimization-based controller synthesis method such as velocity obstacles since the underlying formalism is obtained from duality theory. Additionally, to our knowledge, we are the first to use graph enumeration and KKT conditions to highlight the combinatorial nature of geometric configurations admissible in deadlock. For deadlock resolution, we do not consider additive perturbations since they lack formal guarantees. Instead, we exploit the derived geometric properties to guide the design of a provably correct controller that ensures safety, performance and deadlock resolution.
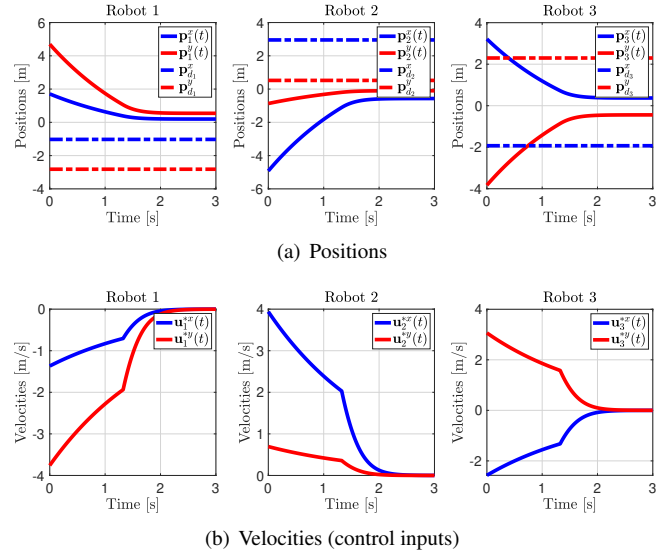
3 Avoidance Control with CBFs: Review

In this section, we review CBF based QPs used for synthesizing controllers that mediate between safety (collision avoidance) and performance (goal-stabilization) for multi-robot systems. We refer the reader to Wang et al. [2017] for a comprehensive treatment on this subject, since our work builds on top of their approach. Assume we have N robots, where each robot follows single-integrator dynamics (we refer the reader to Grover et al. [2019] where we considered double-integrators):

$$\dot{\boldsymbol{p}}_i = \boldsymbol{u}_i. \tag{1}$$

Here $\boldsymbol{p}_i = (x_i, y_i) \in \mathbb{R}^2$ is the position of robot i, $\boldsymbol{u}_i \in \mathbb{R}^2$ is its velocity (*i.e.* control input) and $i \in \{1, 2, \cdots, N\} = [\mathbb{N}]$. In this paper, we ignore actuator limit constraints $|\boldsymbol{u}_i| \leq \alpha_i$ but the reader can refer to our related work Grover et al. [2019] where those constraints were considered. The problem of goal stabilization with collision-avoidance requires that each robot i must reach a goal position $\boldsymbol{p}_{d_i}$ while avoiding collisions with every other robot j, $\forall j \in [\mathbb{N}] \backslash i$. Assume that there is a user-prescribed proportional controller $\hat{\boldsymbol{u}}_i(\boldsymbol{p}_i) = -k_{p_i}(\boldsymbol{p}_i - \boldsymbol{p}_{d_i})$ that generates movement towards goal. Here k_{p_i} is the controller gain for robot i. For posing the collision-free requirement, a pairwise function is formulated that maps the joint state space of i and j to a real-valued safety index *i.e.* $h : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$:

$$h_{ij} = \left\| \Delta \boldsymbol{p}_{ij} \right\|^2 - D_s^2, \tag{2}$$

(a) Positions

(b) Velocities (control inputs)

Figure 2. Positions and velocities of robots falling in deadlock. Note that $\lim_{t \to \infty} p_{x,y} \neq p_{d_{x,y}}$ yet $\lim_{t \to \infty} v_{x,y} = 0$. Simulation results at https://youtu.be/FShai_JESII and experimental demonstration at https://youtu.be/e6eOeuh7Uec.

where D_s is a desired safety margin. Robots i and j are considered to be collision-free or safe if their positions $(\boldsymbol{p}_i, \boldsymbol{p}_j)$ are such that $h_{ij}(\boldsymbol{p}_i, \boldsymbol{p}_j) \geq 0$ (*i.e.* at-least D_s distance apart). We define "pairwise safe set" as the 0-level superset of h_{ij} *i.e.* $\mathcal{C}_{ij} := \{(\boldsymbol{p}_i, \boldsymbol{p}_j) \in \mathbb{R}^4 \mid h_{ij}(\boldsymbol{p}_i, \boldsymbol{p}_j) \geq 0\}$. The boundary of the pairwise safe set is

$$\partial \mathcal{C}_{ij} = \{(\boldsymbol{p}_i, \boldsymbol{p}_j) \in \mathbb{R}^2 | h_{ij}(\boldsymbol{p}_i, \boldsymbol{p}_j) = 0\}. \tag{3}$$

In other words, the system state is on the boundary of the safe set when the robots are grazing each other *i.e.* separated by the safety margin D_s. Assuming that initial positions of robots i, j are collision-free *i.e.* $h_{ij}(\boldsymbol{p}_i(0), \boldsymbol{p}_j(0)) \geq 0$, we would like to synthesize controls $\boldsymbol{u}_i, \boldsymbol{u}_j$ that ensure their future positions are also collision-free *i.e.* $h_{ij}(\boldsymbol{p}_i(t), \boldsymbol{p}_j(t)) \geq 0$, $\forall t > 0$. This can be achieved by ensuring that (Ames et al. [2017]):

$$\frac{dh_{ij}}{dt} \geq -\gamma h_{ij}, \tag{4}$$

where $\gamma > 0$ (a user-defined hyperparameter). Substituting (2) in (4), we get

$$-\Delta \boldsymbol{p}_{ij}^T \Delta \boldsymbol{u}_{ij} \leq \frac{\gamma}{2} h_{ij}. \tag{5}$$

This constraint is distributed on robots i and j as:

$$-\Delta \boldsymbol{p}_{ij}^T \boldsymbol{u}_i \leq \frac{\gamma}{4} h_{ij} \quad \text{and} \quad \Delta \boldsymbol{p}_{ij}^T \boldsymbol{u}_j \leq \frac{\gamma}{4} h_{ij} \tag{6}$$

Therefore, any $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ that satisfy (6) are guaranteed to ensure collision free trajectories for robots i and j in the multirobot system. Since robot i wants to avoid collisions with $N - 1$ robots, there are $N - 1$ collision avoidance constraints. To mediate between safety and goal stabilization objective, a QP is posed that computes a controller closest to the prescribed control $\hat{\boldsymbol{u}}_i(\boldsymbol{p}_i)$ and satisfies the $N - 1$ constraints, as shown:

$$\begin{aligned} \boldsymbol{u}_i^* = \arg\min_{\boldsymbol{u}_i} \quad & \|\boldsymbol{u}_i - \hat{\boldsymbol{u}}_i(\boldsymbol{p}_i)\|^2 \\ \text{subject to} \quad & \boldsymbol{a}_{ij}^T \boldsymbol{u}_i \leq b_{ij} \ \forall j \in [\mathbb{N}] \backslash i \end{aligned} \tag{7}$$

where using (6), we define

$$\boldsymbol{a}_{ij} \coloneqq -\Delta \boldsymbol{p}_{ij}, b_{ij} \coloneqq \frac{\gamma}{4} h_{ij} = \frac{\gamma}{4}(\left\|\Delta \boldsymbol{p}_{ij}\right\|^2 - D_s^2) \quad (8)$$

Each robot i locally solves this QP to determine its $\boldsymbol{u}_i^*$, which guarantees collision avoidance of robot i with $N-1$ robots while encouraging motion towards its goal. While provably safe, this approach does not guarantee that stabilizing to goals will be accomplished because goal stabilization is expressed as a cost function unlike safety which is expressed as a hard constraint. As an example, Fig. 2 shows the result of executing (7) on three robots. Notice from Fig. 2(a) that the positions of robots have converged, but **not** to their respective goals. However, the control inputs from (7) (Fig. 2(b)) have already converged to zero. An experimental realization can be seen at https://youtu.be/e6eOeuh7Uec where three Khepera robots fall in deadlock after using these controllers. Manual intervention is needed to resolve deadlock. Based on these observations, deadlock is defined as (Grover et al. [2020a]):

Definition 1. *Robot i is in deadlock if $\boldsymbol{u}_i^* = 0$ and the prescribed nominal control $\hat{\boldsymbol{u}}_i \neq \boldsymbol{0} \iff \boldsymbol{p}_i \neq \boldsymbol{p}_{d_i}$*

In simpler terms, this definition states that for a robot to be in deadlock, the output from the QP based controller is zero, even though the prescribed controller reports non-zero velocity because the robot is not at its intended destination. We do not analyze "livelock" in this paper because in our simulations we did not observe incidence of oscillatory behavior. Therefore, the focus is on deadlock. In the next section, we examine the closed-loop dynamics of the system *before deadlock* as a function of the robots' initial conditions, goal locations and controller parameters k_{p_i} to characterize what causes deadlock.

4 Before Deadlock

In this section, we analytically investigate the solutions of (7) $\forall i \in \{1, 2, \cdots, N\}$ because these solutions govern the instantaneous dynamics of the robots. However, owing to the fact that (7) is an optimization problem, the controls $\boldsymbol{u}_i^*$ are not available as explicit function of the robots' states. This prohibits straightforward use of tools like Lyapunov theory to analyze long-term behavior of the robots' motion. To circumvent this issue, we use duality theory, specifically KKT conditions of (7) to derive an explicit expression for $\boldsymbol{u}_i^*$ as a function of the robots' positions. These conditions are necessary and sufficient for a global optimum of this QP. The Lagrangian is

$$L(\boldsymbol{u}_i, \boldsymbol{\mu}_i) = \|\boldsymbol{u}_i - \hat{\boldsymbol{u}}_i\|_2^2 + \sum_{j \in [\mathbb{N}] \backslash i} \mu_{ij}(\boldsymbol{a}_{ij}^T \boldsymbol{u}_i - b_{ij})$$

Let $(\boldsymbol{u}_i^*, \boldsymbol{\mu}_i^*)$ be the optimal primal-dual solution to (7). The KKT conditions are Boyd and Vandenberghe [2004]:

1. Stationarity: $\nabla_{\boldsymbol{u}_i} L(\boldsymbol{u}_i, \boldsymbol{\mu}_i)|_{(\boldsymbol{u}_i^*, \boldsymbol{\mu}_i^*)} = 0$

$$\implies \boldsymbol{u}_i^* = \hat{\boldsymbol{u}}_i - \frac{1}{2} \sum_{j \in [\mathbb{N}] \backslash i} \mu_{ij}^* \boldsymbol{a}_{ij}. \quad (9)$$

2. Primal Feasibility

$$\boldsymbol{a}_{ij}^T \boldsymbol{u}_i^* \leq b_{ij} \; \forall j \in [\mathbb{N}] \backslash i \quad (10)$$

3. Dual Feasibility

$$\mu_{ij}^* \geq 0 \; \forall j \in [\mathbb{N}] \backslash i \quad (11)$$

4. Complementary Slackness

$$\mu_{ij}^* \cdot (\boldsymbol{a}_{ij}^T \boldsymbol{u}_i^* - b_{ij}) = 0 \; \forall j \in [\mathbb{N}] \backslash i \quad (12)$$

We define the set of active and inactive constraints as

$$\mathcal{A}(\boldsymbol{u}_i^*) = \{[\mathbb{N}] \backslash i \mid \boldsymbol{a}_{ij}^T \boldsymbol{u}_i^* = b_{ij}\} \quad (13)$$

$$\mathcal{IA}(\boldsymbol{u}_i^*) = \{[\mathbb{N}] \backslash i \mid \boldsymbol{a}_{ij}^T \boldsymbol{u}_i^* < b_{ij}\} \quad (14)$$

These two sets define a partition of the set of non-ego robots. As we will explain in section 5, intuitively, the active set is those set of robots that the ego robot "worries" about for a collision and therefore "actively" repels to stay safe. Using complementary slackness from (12), we deduce

$$\mu_{ij}^* = 0 \; \forall j \in \mathcal{IA}(\boldsymbol{u}_i^*) \quad (15)$$

Therefore, we can restrict the summation in (9) to only the set of active constraints *i.e.*

$$\boldsymbol{u}_i^* = \hat{\boldsymbol{u}}_i - \frac{1}{2} \sum_{j \in \mathcal{A}(\boldsymbol{u}_i^*)} \mu_{ij}^* \boldsymbol{a}_{ij} \quad (16)$$

This equation says that the overall control returned by (7) is the sum of the task-based nominal control $\hat{\boldsymbol{u}}_i$ and the resultant of collision-avoidance velocities from the active robots. While we state this as a fact here, we will prove this in section 5 using a geometric argument. We will use this representation of controls to explore scenarios in which deadlock occurs. Our aim is to characterize the parameters and the geometry of robot arrangements which necessarily result in deadlock instead of characterizing every possible geometric arrangement where deadlock will occur, since this latter problem is combinatorially complex in the number of robots (see section 5). Therefore, to convey the idea, we focus on the simpler cases of two and three robots.

### Closed-loop Analysis for Two Robots

Consider two robots positioned on the line connecting their goals at $t = 0$ as shown in Fig. 3. We will demonstrate that even if the prescribed control of one robot is more aggressive than the other, (7) will necessarily result in deadlock because of the collinearity of initial conditions and goals. Let $\boldsymbol{p}_i(t) \in \mathbb{R}^2$ denote the position of robot $i \in \{1, 2\}$ at time t. At $t = 0$,

$$\boldsymbol{p}_1(0) = (x_0, y_0)$$
$$\boldsymbol{p}_2(0) = \boldsymbol{p}_1(0) + D_{init} \hat{\boldsymbol{e}}_\alpha, \quad (17)$$

where $\hat{\boldsymbol{e}}_\alpha$ is a unit vector oriented at α relative to X_w axis and $D_{init} \in \mathbb{R}^+$ is the initial distance between the robots (see Fig. 3). The desired goal positions are

$$\boldsymbol{p}_{d_1} = \boldsymbol{p}_1(0) + D_{G_1} \hat{\boldsymbol{e}}_\alpha$$
$$\boldsymbol{p}_{d_2} = \boldsymbol{p}_2(0) - D_{G_2} \hat{\boldsymbol{e}}_\alpha. \quad (18)$$

Here D_{G_i} is the distance of i from its goal $\boldsymbol{p}_{d_i}$ at $t = 0$. We assume that $D_{G_1}, D_{G_2} > D_{init}$ to encourage incidence of deadlock. We formally state the deadlock incidence result as follows:
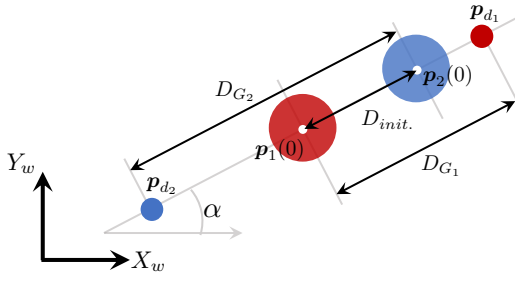
Figure 3. Geometric configuration of two robots and their goals. Notice that everything is collinear along $\hat{e}_\alpha$.

Theorem 1. *Given the initial positions of robots as in (17) and desired goals as in (18), the controls generated by the QP in (7) will cause the robots to fall in deadlock $\forall D_{G_1}, D_{G_2} > D_{init.}$ and $\forall k_{p_1}, k_{p_2} \in \mathbb{R}^+$.*

Proof. We give the sketch of the proof to convey the overall idea and break it down into shorter lemmas. Recall from (16) that the control input on robot i at every time step, depends on the collision avoidance constraints that are active. Therefore, to begin with, we analyze the controls returned by (7) at $t = 0$. In **Lemma 1**, we show that if the initial distance between the robots is greater than a certain critical distance, then the collision avoidance constraints of both robots are inactive, and as a result $u_i^*(0) = \hat{u}_i(0) \forall i$. We then extend this analysis to future time $t > 0$ and show that there exist three sequential phases of motion which are dependent on which robot's collision avoidance constraint is active/inactive:

1. Phase 1 corresponds to the duration in which the collision avoidance constraints of both robots are inactive. We show that this phase culminates in a finite time and the robots move closer to each other during this phase (**Lemma 2**).
2. Next, phase 2 begins where one robot's constraint becomes active while the other's is still inactive. This is due to the heterogeneity $k_{p_1} \neq k_{p_2}$ and $D_{G_1} \neq D_{G_2}$. We show that phase 2 also culminates and the robots move further closer to each other (**Lemma 3**).
3. Finally, phase 3 begins, where both constraints are active. We show that the distance between the robots converges to the safety margin D_s and the robots stop moving (**Lemma 4**), while still away from their goals, thus proving that they have fallen in *deadlock*. □

For the special case of two robots, the control for robot i can be written by adapting (16),

$$u_i^* = \hat{u}_i - \frac{1}{2}\mu_{ij}^* a_{ij}, \qquad (19)$$

where $j \neq i$ and $i \in \{1, 2\}$. In (15), we showed that the value of Lagrange multiplier μ_{ij} depends on whether the collision avoidance constraint $a_{ij}^T u_i - b_{ij} \leq 0$ is active (=) or inactive (<) at $u_i = u_i^*$. Thus, we have a check flag for evaluating the feasibility of a candidate control $\tilde{u}_i$:

$$f_{ij}(\tilde{u}_i) = a_{ij}^T \tilde{u}_i - b_{ij}, \qquad (20)$$

$\forall i \in \{1, 2\}$ and $j \neq i$. If for a candidate control $\tilde{u}_i$, $f_{ij}(\tilde{u}_i) > 0$, then $\tilde{u}_i$ violates robot $i's$ constraint with j

in (7). Let us check this flag for the nominal control $\hat{u}_i$. Suppose that $f_{ij}(\hat{u}_i) < 0$ *i.e.* $\hat{u}_i$ ensures that the constraint of i with j is feasible. Then, looking at (7), notice that its optimum will be attained at $u_i^* = \hat{u}_i$ itself with zero cost (as long as there are only two robots in the system). This is because $\hat{u}_i$ by itself satisfies the constraint and the cost function penalizes deviation from $\hat{u}_i$. Further, $f_{ij}(u_i^*) = f_{ij}(\hat{u}_i) < 0$ implies that the constraint of i with j is inactive.

Let us look at the robots geometrically to see when this happens. Define critical distance for robot i as (to be derived in **Lemma 1**):

$$\beta_+^i = \frac{2D_{G_i}k_{p_i}}{\gamma} + \sqrt{\left(\frac{2D_{G_i}k_{p_i}}{\gamma}\right)^2 + D_s^2} \qquad (21)$$

$\forall i \in \{1, 2\}$. We now show that if the initial distance between the robots i and j *i.e.* $D_{init.}$, is greater than i's critical distance β_+^i, then $i's$ constraint with j is inactive at $t = 0$.

Lemma 1. *If at $t = 0$, $D_{init} > \beta_+^i$, then $u_i^*(0) = \hat{u}_i(0)$.*

Intuition: This result says that if at $t = 0$, robot i is far from j (specifically at-least β_+^i distance away), then i can simply use it's prescribed nominal control $\hat{u}_i$ at $t = 0$, and not worry about a collision with j at-least until the next instant.

Proof. See appendix (A.1).

Assumption 1. *WLOG assume that $D_{init} > \beta_+^1 > \beta_+^2$*

From assumption (1), and the result of **Lemma 1**, it follows that the collision avoidance constraints of both robots are inactive at $t = 0$. Thus, $u_i^*(0) = \hat{u}_i(0)$ for $i = \{1, 2\}$. If we give these velocities to robots for a small time Δt, the distance between the robots and the critical distance will change. So at the next instant, we will compare the *updated inter-robot distance* with the *updated critical distances* to decide whether $u_i^*(\Delta t) = \hat{u}_i(\Delta t)$ and so on. Therefore, one can assume that these constraints remain inactive for some finite time, which is precisely the duration of phase 1.

Phase 1: We define phase 1 as the period in which the constraints of both robots are inactive *i.e.* both their flags are strictly negative when evaluated using their nominal controllers. In other words,

$$f_{12}(\hat{u}_1(t)) = a_{12}^T(t)\hat{u}_1(t) - b_{12}(t) < 0 \qquad (22)$$
$$f_{21}(\hat{u}_2(t)) = a_{21}^T(t)\hat{u}_2(t) - b_{21}(t) < 0 \qquad (23)$$

Mathematically, the duration of phase 1 is given by,

$$t_1 := \sup_{t>0}\{t \,|\, f_{12}(\hat{u}_1(t)) < 0, f_{21}(\hat{u}_2(t)) < 0\}, \qquad (24)$$

i.e. the maximum time until which the nominal controllers of both robots remain feasible. So until t_1, we have

$$\dot{p}_1 = u_1^* = \hat{u}_1 \text{ and } \dot{p}_2 = u_2^* = \hat{u}_2 \qquad (25)$$

Now given the initial positions and desired goals as in Fig. 3, it cannot be the case that both robots continue to use their nominal controllers indefinitely to move to their goals because that would result in a collision. Thus, a time will come when using the nominal control is no longer feasible for at-least one robot. This is precisely the time t_1 defined in Def. (24) and we prove its existence in the next lemma.

Lemma 2. *∃ a finite time t_1 as described in Def. (24), until which the collision avoidance constraints of both robots are simultaneously inactive.*

Intuition: This result states that for sometime, both robots may continue to use their nominal controllers and move without "worrying" about collisions with the other. Nevertheless, eventually at t_1, the inter-robot distance will fall below a critical threshold for one robot, at which point, its collision avoidance constraint will get activated. For example, in Fig. 4, the dark green curve (the inter-robot distance) begins to fall below the blue curve (the critical distance of robot one (26)) at t_1. Thus, robot one can no longer use its nominal control starting from t_1.

Proof. See appendix A.2.

The critical distance for robot one (blue curve in Fig. 4) is derived in the appendix A.2 and is given by:

$$\beta_+^1(t) = \frac{2D_{G_1}k_{p_1}e^{-k_{p_1}t}}{\gamma} + \sqrt{\left(\frac{2D_{G_1}k_{p_1}e^{-k_{p_1}t}}{\gamma}\right)^2 + D_s^2}$$

(26)

Thus, phase 1 ends at t_1 marking the beginning of phase 2.

Phase 2: We define this phase as the period in which the constraint of robot one has become active *i.e.* $\boldsymbol{a}_{12}^T \boldsymbol{u}_1^* = b_{12}$ while that of robot two is inactive *i.e.* $\boldsymbol{a}_{21}^T \boldsymbol{u}_2^* < b_{21}$. So for $t \geq t_1$, the dynamics of the robots are

$$\dot{\boldsymbol{p}}_1 = \boldsymbol{u}_1^* = \hat{\boldsymbol{u}}_1 - \frac{1}{2}\mu_{12}\boldsymbol{a}_{12} \text{ , where } \mu_{12} = 2\frac{\boldsymbol{a}_{12}^T\hat{\boldsymbol{u}}_1 - b_{12}}{\|\boldsymbol{a}_{12}\|^2}$$

$$\dot{\boldsymbol{p}}_2 = \boldsymbol{u}_2^* = \hat{\boldsymbol{u}}_2$$

Here we derived $\boldsymbol{u}_1^*$ using (19). With some calculation, we can show that the components of velocities in $\boldsymbol{u}_{\{1,2\}}^*$ perpendicular to $\hat{\boldsymbol{e}}_\alpha$ vanish. Now, owing to the inevitability of a collision should robot two choose to continue using $\hat{\boldsymbol{u}}_2$ forever, eventually, the constraint of robot two will also switch from inactive to active (A.3). Therefore, similar to the definition of t_1, we define the duration of phase 2 until the time for which the constraint of robot two stays inactive *i.e.*

$$t_2 := \sup_{t > t_1}\{t \mid f_{21}(\hat{\boldsymbol{u}}_2(t)) < 0\},$$

(27)

Lemma 3. *∃ a finite time t_2 as in Def. (27), until which the constraint of robot two stays inactive.*

Intuition: Eventually there will come a time t_2 when the inter-robot distance will fall just below the critical threshold of robot two as well, just the way it did for robot one at t_1. So at this time, robot two's collision avoidance constraint will become active (see Fig. 4 where the dark green curve (the inter-robot distance) intersects the yellow curve (the critical distance of robot two (28)).

Proof. See appendix A.3.

The critical distance for robot two (yellow curve in Fig. 4) is defined similar to (26) as:

$$\beta_+^2(t) = \frac{2D_{G_2}k_{p_2}e^{-k_{p_2}t}}{\gamma} + \sqrt{\left(\frac{2D_{G_2}k_{p_2}e^{-k_{p_2}t}}{\gamma}\right)^2 + D_s^2}$$

(28)
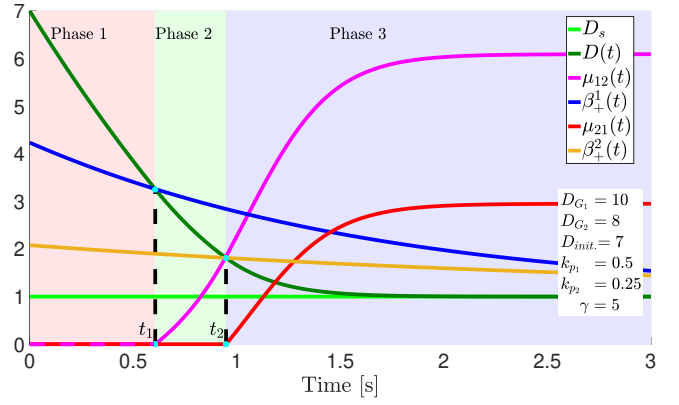
Figure 4. Simulation results for two-robot deadlock. Simulation parameters shown in bottom right panel. Note at $t = t_1$, $D(t) = \beta_+^1(t)$ and $\mu_{12}(t)$ switches on, similarly at $t = t_2$, $\mu_{21}(t)$ switches on. Finally, $D(t) \longrightarrow D_s$.

Phase 3: This phase begins at t_2 and in this phase, the constraints of both robots are active *i.e.* $\boldsymbol{a}_{12}^T\boldsymbol{u}_1^* = b_{12}$ and $\boldsymbol{a}_{21}^T\boldsymbol{u}_2^* = b_{21}$. So for $t \geq t_2$, the dynamics of the robots are

$$\dot{\boldsymbol{p}}_1 = \boldsymbol{u}_1^* = \hat{\boldsymbol{u}}_1 - \frac{1}{2}\mu_{12}\boldsymbol{a}_{12} \text{ , where } \mu_{12} = 2\frac{\boldsymbol{a}_{12}^T\hat{\boldsymbol{u}}_1 - b_{12}}{\|\boldsymbol{a}_{12}\|^2}$$

$$\dot{\boldsymbol{p}}_2 = \boldsymbol{u}_2^* = \hat{\boldsymbol{u}}_2 - \frac{1}{2}\mu_{21}\boldsymbol{a}_{21} \text{ , where } \mu_{21} = 2\frac{\boldsymbol{a}_{21}^T\hat{\boldsymbol{u}}_2 - b_{21}}{\|\boldsymbol{a}_{21}\|^2}.$$

We can show that the velocities of both robots *i.e.* $\boldsymbol{u}_{\{1,2\}}^*$ have no component perpendicular to $\hat{\boldsymbol{e}}_\alpha$. Recall that $\boldsymbol{a}_{12} = -\boldsymbol{a}_{21} = \Delta\boldsymbol{p}_{21}$ is parallel to $\hat{\boldsymbol{e}}_\alpha$. Additionally, $b_{12} = b_{21} = \frac{\gamma}{4}(\|\Delta\boldsymbol{p}_{21}\| - D_s^2)$. Using these, we can simplify the dynamics of the robots to

$$\dot{\boldsymbol{p}}_1 = +\gamma\frac{\|\Delta\boldsymbol{p}_{21}\|^2 - D_s^2}{4\|\Delta\boldsymbol{p}_{21}\|^2}\Delta\boldsymbol{p}_{21}$$

(29)

$$\dot{\boldsymbol{p}}_2 = -\gamma\frac{\|\Delta\boldsymbol{p}_{21}\|^2 - D_s^2}{4\|\Delta\boldsymbol{p}_{21}\|^2}\Delta\boldsymbol{p}_{21}$$

(30)

$$\implies \dot{\Delta\boldsymbol{p}}_{21} = -\gamma\frac{\|\Delta\boldsymbol{p}_{21}\|^2 - D_s^2}{2\|\Delta\boldsymbol{p}_{21}\|^2}\Delta\boldsymbol{p}_{21}$$

(31)

Given this inter-robot dynamics, we show in A.4 that the asymptotic solution of (31) proves that $\|\Delta\boldsymbol{p}_{21}(t)\|$ *i.e.* the inter-robot distance converges to D_s and additionally, both robots come to a halt while still being away from their goals. This establishes the incidence of deadlock.

Lemma 4. *The distance between robots converges to the safety margin D_s at which point they stop moving and fall in deadlock.*

Proof. See appendix A.4.

Takeaway: We have demonstrated that two robots fall in deadlock given initial conditions (17) and goals (18). This is because the *geometric arrangement of initial positions and goals prevents both $\hat{\boldsymbol{u}}$ and the QP to output velocities perpendicular to the $\hat{\boldsymbol{e}}_\alpha$ direction* in all the three phases. With the control always along $\boldsymbol{u}_\parallel \parallel \hat{\boldsymbol{e}}_\alpha$, the only way to reach the goals would involve intersection of the robots *i.e.* collisions. This is why, the QP based controller decides to prevent collisions from happening by forcing the robots to stall, resulting in deadlock.

Closed-Loop Analysis for Three-Robots

We describe another scenario prone to deadlock, which consists of three robots positioned on the vertices of an equilateral triangle, and required to stabilize to their respective antipodal positions. For simplicity, we assume that all prescribed controller gains are *identical*, as are the distances of the robots' initial positions to their goals. Let $\boldsymbol{p}_i(t) \in \mathbb{R}^2$ be the position of robot $i \in \{1, 2, 3\}$ given as

$$\boldsymbol{p}_1(0) = (x_0, y_0)$$
$$\boldsymbol{p}_2(0) = \boldsymbol{p}_1(0) + D_{init}\hat{\boldsymbol{e}}_\alpha$$
$$\boldsymbol{p}_3(0) = \boldsymbol{p}_2(0) + D_{init}\hat{\boldsymbol{e}}_{\alpha+\frac{2\pi}{3}} \quad (32)$$

Assume that the goal of each robot is diametrically opposite to its initial position *i.e.*

$$\boldsymbol{p}_{d_i} = \boldsymbol{p}_i(0) + \frac{\sqrt{3}D_G}{D_{init.}}(\boldsymbol{c} - \boldsymbol{p}_i(0)) \quad (33)$$

for $i \in \{1, 2, 3\}$ where $\boldsymbol{c} = \frac{1}{3}\sum_{i=1}^{3}\boldsymbol{p}_i(0)$ is the centroid of the equilateral triangle formed by the initial positions and D_G is the distance of each robot from its goal. This scenario is prone to deadlock, because of (1) the geometric symmetry in the initial positions of robots and goals and (2) identical controller gains.

Assumption 2. $\sqrt{3}D_G > D_{init.}$

This assumption will be needed in **Lemma 8** to establish the inevitability of deadlock. We now formally state the deadlock incidence result as follows.

Theorem 2. *Given the initial positions of robots as in* (32) *and desired goals as in* (33), *the controls generated by* (7) *will cause the robots to fall in deadlock.*

Proof. Our proof draws upon the ideas from deadlock in the two robot case. The overall sketch is as follows:

1. First, we analyze the controls returned by the QP at $t = 0$. We show in **Lemma 5** that if the initial distance between robots is greater than a certain critical distance, then at $t = 0$, all collision avoidance constraints of all robots are inactive, therefore $\boldsymbol{u}_i^*(0) = \hat{\boldsymbol{u}}_i(0) \, \forall i \in \{1, 2, 3\}$.

2. Next, we show that the future motion of robots can be broken into two successive phases. In phase 1, the collision avoidance constraints of all robots stay inactive, so they use $\hat{\boldsymbol{u}}_i(t) \, \forall i \in \{1, 2, 3\}$. We prove in **Lemma 6** that using $\hat{\boldsymbol{u}}_i(t)$, the robots move on the vertices of an equilateral triangle. Next, in **Lemma 7**, we demonstrate that there exists a finite time when phase 1 culminates (see Fig. 5). This time is identical for all robots because of the symmetry that follows from **Lemma 6**.

3. Next, phase 2 begins during which all constraints become active. We show in **Lemma 8** that the distance between robots converges to D_s, their velocities converge to zero while the robots are still away from their goals, thus establishing *deadlock*. $\square$

Recall from (16) that the control for robot i is

$$\boldsymbol{u}_i^* = \hat{\boldsymbol{u}}_i - \frac{1}{2}\sum_{j \in \{1,2,3\}\setminus i}\mu_{ij}\boldsymbol{a}_{ij} \quad (34)$$
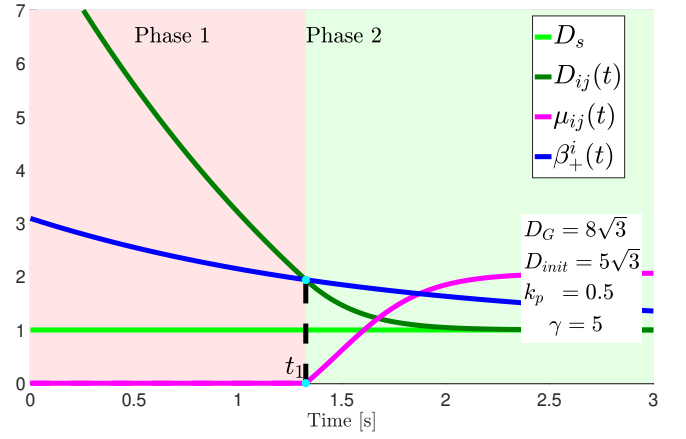
Figure 5. Simulation results for three-robot deadlock. Simulation parameters shown in bottom right panel. Note at $t = t_1$, $D_{ij}(t) = \beta_+^i(t)$ and $\mu_{ij}(t)$ switches on. Finally, $D(t) \longrightarrow D_s$.

Here, the value of μ_{ij} depends on whether the constraint $\boldsymbol{a}_{ij}^T\boldsymbol{u}_i^* - b_{ij} \leq 0$ is active ($= 0$) or inactive (< 0). So like the two robot case, we check the value of the flag $f_{ij}(\tilde{\boldsymbol{u}}_i) = \boldsymbol{a}_{ij}^T\tilde{\boldsymbol{u}}_i - b_{ij}$ for the nominal controller *i.e.* $\tilde{\boldsymbol{u}}_i = \hat{\boldsymbol{u}}_i$. We show that if the initial distance between *any two* robots is greater than the critical distance β_+ (35), then at $t = 0$, all robots have all their constraints inactive. Consequently, the optimal control returned by (7) will be $\hat{\boldsymbol{u}}_i(0)$. The critical distance is defined as (derived in A.5):

$$\beta_+ = \frac{\sqrt{3}D_G k_p}{\gamma} + \sqrt{\left(\frac{\sqrt{3}D_G k_p}{\gamma}\right)^2 + D_s^2} \quad (35)$$

Lemma 5. *If* $D_{init.} > \beta_+$, *then* $\boldsymbol{u}_i^*(0) = \hat{\boldsymbol{u}}_i(0) \, \forall i$

Intuition: This result states that, if at $t = 0$, i is away from j by at-least β_+, then i can simply use $\hat{\boldsymbol{u}}_i(0)$ and not "worry" about a collision with j at-least until the next instant. This holds for all robots because β^+ is identical for them all due to the assumed symmetry and their identical controller gains.

Proof. See appendix A.5.

Assumption 3. *WLOG assume at* $t = 0$, $D_{init.} > \beta_+$.

From assumption (3) and **Lemma 5**, it follows that at $t = 0$, **all** collision avoidance constraints of **all** robots are inactive. Therefore, it is safe to assume that these constraints also remain inactive for a finite duration after $t = 0$, which is precisely the duration of phase 1.

Phase 1: In this phase, the robots use their nominal controls. Thus, their dynamics are given by:

$$\dot{\boldsymbol{p}}_i = \boldsymbol{u}_i^* = \hat{\boldsymbol{u}}_i = -k_p(\boldsymbol{p}_i - \boldsymbol{p}_{d_i})$$
$$\implies \boldsymbol{p}_i(t) = e^{-k_p t}\boldsymbol{p}_i(0) + (1 - e^{-k_p t})\boldsymbol{p}_{d_i} \quad (36)$$

Next we show that given the choice of symmetric initial and goal positions as in (32) and (33) respectively and identical controller gains, the robots retain this symmetry and move along the vertices of an equilateral triangle.

Lemma 6. *All robots continue to move in an equilateral triangular configuration with centroid fixed at* $\boldsymbol{c}(0)$

Intuition: Since the initial positions and goals were symmetric, and since all robots have identical controller gains, they move by the same distance per time step. So the symmetry is preserved every time which is why they retain the equilateral triangular configuration. https://bit.ly/2VnDXdb conveys this idea.

Proof. See appendix A.6.

Next we show that eventually, there comes a time when it is no longer feasible to use the nominal control $\hat{\boldsymbol{u}}_i(t) \, \forall i$ at which point all the constraints of all robots become active. This again follows from the inevitability of collisions should all three robots continue using $\hat{\boldsymbol{u}}_i$ indefinitely.

Lemma 7. *There exists a time t_{ij} when the collision avoidance constraint of robot i with robot j becomes active. Furthermore, t_{ij} is identical $\forall i \in \{1, 2, 3\}, \, \forall j \in \{1, 2, 3\} \backslash i$.*

Intuition: Eventually, the inter-robot distance (same for any pair) falls below the common critical distance (see Fig. 5 where the dark green curve (the inter-robot distance) intersects the blue curve (the critical distance)). Therefore, it is no longer feasible to use the nominal control. This occurs simultaneously for all robots due to the geometric symmetry and identical controller parameters. Let this common time be t_1 as shown in Fig 5.

Proof. See appendix A.7.

Phase 2: For $t \geq t_1$, all the constraints of all robots have become active. Thus, the dynamics of the robots are:

$$\dot{\boldsymbol{p}}_i = \boldsymbol{u}_i^* = \hat{\boldsymbol{u}}_i - \frac{1}{2} \sum_{j \in \{1,2,3\} \backslash i} \mu_{ij} \boldsymbol{a}_{ij} \qquad (37)$$

where the expressions for $\mu_{ij} \neq 0$ can be derived by using the activeness of the constraints *i.e.* $\boldsymbol{a}_{ij}^T \boldsymbol{u}_i^* = b_{ij}$. Following arguments in phase three for the two robot case, we can show that the positions of the robots can be written as:

$$\begin{aligned}
\boldsymbol{p}_1(t) &= \eta(t)\boldsymbol{p}_1(0) + (1 - \eta(t))\boldsymbol{p}_{d_1} \\
\boldsymbol{p}_2(t) &= \boldsymbol{p}_1(t) + D(t)\hat{\boldsymbol{e}}_\alpha \\
\boldsymbol{p}_3(t) &= \boldsymbol{p}_2(t) + D(t)\hat{\boldsymbol{e}}_{\alpha + \frac{2\pi}{3}}
\end{aligned} \qquad (38)$$

for $t \geq t_1$ for some function $\eta(t)$ with $\eta(t_1) = e^{-k_p t_1}$. Here $D(t)$ is the inter-robot distance which is same for any two pairs. Using (37), (38), and the definition of $\boldsymbol{a}_{ij}, b_{ij}$ from (8), the dynamics of robot i can be simplified to:

$$\dot{\boldsymbol{p}}_i = \sqrt{3}\gamma \frac{D^2(t) - D_s^2}{6D(t)} \hat{\boldsymbol{e}}_{\alpha + (4i-3)\frac{\pi}{6}} \qquad (39)$$

$$\implies \dot{\Delta \boldsymbol{p}}_{12} = \gamma \frac{D^2(t) - D_s^2}{2D(t)} \hat{\boldsymbol{e}}_\alpha \qquad (40)$$

Given this inter-robot dynamics, we show in A.8 that the asymptotic solution of (40) proves that $\|\Delta \boldsymbol{p}_{ij}(t)\|$ *i.e.* the inter-robot distance converges to D_s and additionally, all three robots come to a halt while still being away from their goals. This establishes the incidence of deadlock.

Lemma 8. *The distance between the robots converges to D_s and the robots stop moving and fall in deadlock.*

Proof. See appendix A.8.

Takeaway: Deadlock happens here because each robot's velocity is always pointing towards its goal and not perpendicular to it. In phase 1, this velocity is along $\hat{\boldsymbol{u}}$ (which by definition is towards the goal). Likewise, we can show that the same holds true in phase 2. Additionally, all robots have identical speed at any given time as well. Hence, collisions would be inevitable at the centroid. Therefore, CBF-QPs cause the robots to stall which results in deadlock.

5   During Deadlock

In the previous section, we analyzed cases of two and three robots and showed that geometric symmetry in the initial positions and goals results in deadlock. In this section, we assume that the system is already in deadlock and infer geometric properties of robot configurations that are valid in deadlock. The analysis in this section is general for N robots, and not restricted to two or three robots. Recall that we defined deadlock as

Definition 2. *Robot i is in deadlock if $\boldsymbol{u}_i^* = 0$ and $\boldsymbol{p}_i \neq \boldsymbol{p}_{d_i}$*

We reformulate the KKT conditions (9)-(12) using this definition to derive geometric properties of the system in deadlock. The aim is to show that these properties are indeed consistent with what one would expect to intuitively hold true when the robots are in deadlock.

Reformulating KKT Conditions for Deadlock and their Mechanics Interpretation:

Primal Feasibility: Recall from (10) that $\boldsymbol{a}_{ij}^T \boldsymbol{u}_i^* \leq b_{ij} \, \forall j \in \{1, 2, \cdots, N\} \backslash i$. Since in deadlock, $\boldsymbol{u}_i^* = \boldsymbol{0}$, this means

$$0 \leq b_{ij} \, \forall j \in [\mathbb{N}] \backslash i \qquad (41)$$

From (8), $b_{ij} = \frac{\gamma}{4} h_{ij} = \frac{\gamma}{4} \left(\|\Delta \boldsymbol{p}_{ij}\|^2 - D_s^2 \right)$ which using (41) implies

$$0 \leq h_{ij} \iff D_s \leq \|\Delta \boldsymbol{p}_{ij}\| \, \forall j \in [\mathbb{N}] \backslash i \qquad (42)$$

This means that robot i, when in deadlock, is at-least D_s distance away from every other robot and therefore safe.

Dual Feasibility: From (11) recall that,

$$\mu_{ij}^* \geq 0 \, \forall j \in [\mathbb{N}] \backslash i \qquad (43)$$

Complementary Slackness: From (12), recall that $\mu_{ij}^* \cdot (\boldsymbol{a}_{ij}^T \boldsymbol{u}_i^* - b_{ij}) = 0 \, \forall j \in [\mathbb{N}] \backslash i$. Substituting $\boldsymbol{u}_i^* = \boldsymbol{0}$ and $b_{ij} = \frac{\gamma}{4} h_{ij} = \frac{\gamma}{4} \left(\|\Delta \boldsymbol{p}_{ij}\|^2 - D_s^2 \right)$, we get

$$\mu_{ij}^* \cdot \left(\|\Delta \boldsymbol{p}_{ij}\|^2 - D_s^2 \right) = 0 \, \forall j \in [\mathbb{N}] \backslash i \qquad (44)$$

Thus, using (43), (44) and (13), the constraints that are active and inactive for robot i in deadlock are

$$\begin{aligned}
\mathcal{A}(\boldsymbol{0}) &= \{j \in [\mathbb{N}] \backslash i \mid \|\Delta \boldsymbol{p}_{ij}\| = D_s \iff h_{ij} = 0\} \\
\mathcal{IA}(\boldsymbol{0}) &= \{j \in [\mathbb{N}] \backslash i \mid \|\Delta \boldsymbol{p}_{ij}\| > D_s\}
\end{aligned} \qquad (45)$$

Thus, active constraints refer to those robots that are exactly D_s away from i. Alternatively, these are the robots j whose positions are such that $h_{ij} = 0$ *i.e.* robot i's safety is at the verge of being compromised thanks to these robots. In the next lemma, we show that if i is in deadlock, then it can never be the case that there are no active robots.
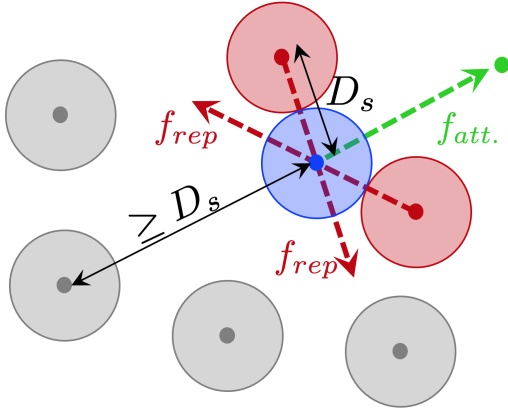
Figure 6. An illustration of the mechanics interpretation of KKT conditions. The ego robot is shown in blue and its desired goal location in green. The green arrow represents the force attracting the ego robot towards its goal. The red robots are the ones in the active set since they are separated by exactly D_s, therefore, they repel the ego robot. The robots shown in gray represent the inactive constraints.

Lemma 9. $\mathcal{A}(\mathbf{0})$ *is non-empty.*

Intuition: When i is in deadlock, $\mathcal{A}(\mathbf{0})$ is the set of robots preventing it from reaching its goal. Thus, this set cannot be empty as other wise the robot would be able to move freely to reach its goal and hence not be in deadlock.

Proof. We prove this by contradiction. Suppose that $\mathcal{A}(\mathbf{0})$ is empty or said another way $\mathcal{IA}(\mathbf{0}) = [\mathbb{N}]\backslash i$. This means that $\mu_{ij}^* = 0 \ \forall j \in [\mathbb{N}]\backslash i$ and hence $\hat{\boldsymbol{u}}_i = \boldsymbol{u}_i^*$ from stationarity (9). Furthermore, since i is in deadlock, from Def. (2) we have that $\boldsymbol{p}_i \neq \boldsymbol{p}_{d_i} \iff \hat{\boldsymbol{u}}_i \neq \mathbf{0}$ since $\hat{\boldsymbol{u}}_i = -k_{p_i}(\boldsymbol{p}_i - \boldsymbol{p}_{d_i})$ However, $\boldsymbol{u}_i^* = \mathbf{0}$ in deadlock. This means that $\boldsymbol{u}_i^* \neq \hat{\boldsymbol{u}}_i$ in deadlock, giving rise to the contradiction. $\qquad \square$

Furthermore, we can interpret the multipliers μ_{ij}^* as spring-constants of springs connecting i to $\mathcal{A}(\mathbf{0})$. This gives rise to the force-equilibrium interpretation of stationarity:

Stationarity: Using (16) and $\boldsymbol{u}_i^* = \mathbf{0}$, we get

$$\hat{\boldsymbol{u}}_i - \frac{1}{2} \sum_{j \in \mathcal{A}(\mathbf{0})} \mu_{ij}^* \boldsymbol{a}_{ij} = \mathbf{0}$$

$$\iff -k_{p_i}(\boldsymbol{p} - \boldsymbol{p}_{d_i}) + \frac{1}{2} \sum_{j \in \mathcal{A}(\mathbf{0})} \mu_{ij}^* \Delta\boldsymbol{p}_{ij} = \mathbf{0} \quad (46)$$

In this equation, the first term $\hat{\boldsymbol{u}}_i = -k_{p_i}(\boldsymbol{p} - \boldsymbol{p}_{d_i})$ represents an attractive force pulling the robot i towards its goal $\boldsymbol{p}_{d_i}$. The second term $+\frac{1}{2}\sum_{j \in \mathcal{A}(\boldsymbol{u}_i^*)} \mu_{ij}^* \Delta\boldsymbol{p}_{ij}$ represents the resultant repulsive force from active robots *i.e.* ones that are D_s away from i. Thus, deadlock occurs when the attractive force to the goal is balanced by the net repulsion from active robots.

Based on these conditions, we motivate a set-theoretic interpretation of deadlock. Let the state of the ego robot be $\boldsymbol{p}_i \in \mathbb{R}^2$ and let the states of the remaining $N - 1$ robots be denoted as $\boldsymbol{p}_{-i} = \{\boldsymbol{p}_j\} \ \forall j \in [\mathbb{N}]\backslash i$. Define $\mathcal{D}_i$ as:

$$\mathcal{D}_i = \{\boldsymbol{p}_i \in \mathbb{R}^2 \mid \boldsymbol{u}_i^*(\boldsymbol{p}_i, \boldsymbol{p}_{-i}) = \mathbf{0}, \boldsymbol{p}_i \neq \boldsymbol{p}_{d_i}\} \quad (47)$$

The set $\mathcal{D}_i$ is defined as those states of robot i which satisfy the criteria of i being in deadlock. We restate the $\boldsymbol{u}_i^*(\boldsymbol{p}_i, \boldsymbol{p}_{-i}) = \mathbf{0}$ criterion in this definition with the reformulated KKT conditions described above:

$$\mathcal{D}_i = \left\{ \boldsymbol{p}_i \in \mathbb{R}^2 \middle| \|\Delta\boldsymbol{p}_{ij}\| \geq D_s, \mu_{ij}^* \geq 0 \ \forall j \in [\mathbb{N}]\backslash i, \right.$$
$$\left. \hat{\boldsymbol{u}}_i(\boldsymbol{p}_i) = \frac{1}{2} \sum_{j \in \mathcal{A}(\mathbf{0})} \mu_{ij}^* \boldsymbol{a}_{ij}, \boldsymbol{p}_i \neq \boldsymbol{p}_{d_i} \right\} \quad (48)$$

The motivation behind stating this definition is to interpret deadlock as a bonafide set in the state space of the ego robot and derive a control strategy that makes the robot evade/exit this set. Building on this definition of robot i's deadlock set, we motivate *system deadlock* to be the joint states of all robots for which all of them are in deadlock:

$$\mathcal{D}_{system} = \left\{ (\boldsymbol{p}_1, \cdots, \boldsymbol{p}_N) \in \mathbb{R}^{2N} \middle| \boldsymbol{p}_i \in \mathcal{D}_i, i \in \right.$$
$$\left. \{1, \cdots, N\} \right\} = \prod_{i=1}^{N} \mathcal{D}_i \quad (49)$$

In the rest of the paper, we will focus on system deadlock. This is because the case where only a subset of robots are in deadlock can be decomposed into subproblems where a subset is in *system deadlock* and the remaining robots are free to move. The next section focuses on complexity analysis of *system deadlock*.

Graph Enumeration based Complexity Analysis

The Lagrange multipliers μ_{ij}^* are in general, a nonlinear function of the states of the robots. Their values depend on which constraints are active/inactive. An active constraint will in-turn determine the set of possible geometric configurations that the robots can take when they are in deadlock and this in turn will guide the design of our deadlock resolution algorithm. Therefore, we are interested in deriving all possible combinations of active/inactive constraints that the robots can assume once in deadlock. But first we derive upper and lower bounds for the number of valid configurations in *system deadlock*.

We can interpret an active constraint between robots i and j as an undirected edge between vertices i and j in a graph formed by N labeled vertices, where each vertex represents a robot. The following property (which follows from symmetry) allows the edges to be undirected.

Lemma 10. *If robot i and j are both in deadlock and i's constraint with j is active (inactive), then j's constraint with i is also active (inactive).*

Intuition: If i and j are in deadlock and i perceives that j is at the verge of violating i's safety, then j also perceives that i is at the verge of violating its safety.

Proof. Since i is in deadlock and its constraint with j is active, we have $\boldsymbol{a}_{ij}^T \boldsymbol{u}_i^* = b_{ij} = 0 \iff b_{ji} = 0$, since from, (8) we have that $b_{ij} = b_{ji}$. Since j is in deadlock, $\boldsymbol{a}_{ji}^T \boldsymbol{u}_j^* = 0 \implies b_{ji} = \boldsymbol{a}_{ji}^T \boldsymbol{u}_j^*$. This means j's constraint with i is also active. $\qquad \square$

Upper Bound Given N vertices, there are $^NC_2$ distinct pairs of edges possible. The overall system can have any subset of those edges. Since a set with $^NC_2$ members has $2^{^NC_2}$ subsets, we conclude that there are $2^{^NC_2}$ possible graphs. In other words, given N robots, the number of configurations that are admissible in deadlock is $2^{^NC_2}$. However, this number is an upper bound because it includes cases where a given vertex can be disconnected from all other vertices, which is not valid in *system deadlock* because $\mathcal{A}(\mathbf{0})$ cannot be empty (**Lemma** 9).

Lower Bound Since $\mathcal{A}(\mathbf{0})$ is non-empty, each vertex in the graph will have at-least one edge *i.e.* each robot will have at-least one constraint active with some other robot. From this observation, it follows that the set of graphs valid in *system deadlock* is a superset of connected simple graphs that can be formed by N labeled vertices. It is also possible that some simply connected graphs may not be feasible due to restrictions imposed by Euclidean geometry.

Graphs that are (a) simple and connected, (b) have N labelled vertices (since each robot has an ID), (c) are embedded in $\mathbb{R}^2$ (environment is planar), (d) have Euclidean distance between connected vertices equal to D_s, (e) that between unconnected vertices greater than D_s, and (f) have one or two edges per vertex (possibly more), necessarily represent admissible geometric configurations of robots in *system deadlock*. This follows because every graph satisfying qualifiers (a)-(f) can be mapped to a geometric arrangement of N robots in deadlock. A lower bound for graphs satisfying all qualifiers (a)-(f) can be shown to be $0.5(N+1)(N-1)!$ as follows ($N \geq 3$).

Consider a cyclic graph whose each node is the vertex of an N regular polygon with side D_s. Such a graph necessarily satisfies (a)-(f). Re-arrangements of its vertices gives rise to $0.5(N-1)!$ graphs. Likewise, a graph with nodes along an open chain also satisfies (a)-(f), and gives $0.5N!$ rearrangements. Thus, the total is $0.5(N-1)! + 0.5N! = 0.5(N+1)(N-1)!$. It is well known that factorial overtakes exponential, thus highlighting the increase in the number of geometric configurations. Our MATLAB simulations show that the exact number of configurations for $N = \{1, 2, 3, 4\}$ are $\{1, 1, 4, 18\}$ whereas our bound gives $\{1, 1, 4, 15\}$. This simulation demonstrates the explosion in possible geometric configurations that are admissible in *system deadlock* with increasing number of robots. Therefore for further analysis, we will restrict to the case of two and three robots.

Characteristics of Two-Robot Deadlock

In this section, we give some specific properties of a two-robot system in deadlock which we will use later to resolve deadlocks in two robots. We show that these properties hold true regardless of the choice of initial conditions.

In the two-robot system, an individual robot by itself cannot be in deadlock *i.e.* either both robots are in deadlock or neither. This is because the sole collision avoidance constraint is symmetric due to **Lemma** 10. From **Lemma** 9, each robot perceives the other robot as active. Define the

(a) Two Robot Equilibrium   (b) Three Robot Equilibrium
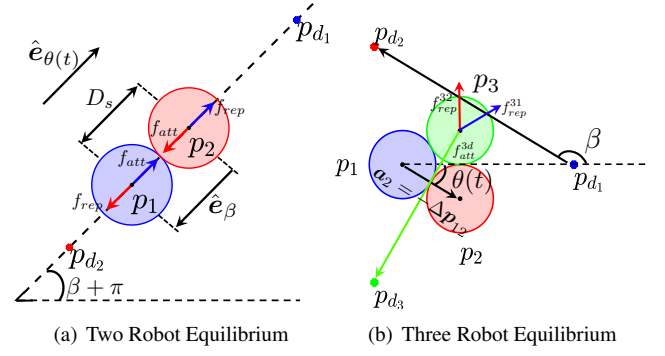
Figure 7. Force Equilibrium in Deadlock

system deadlock set $\mathcal{D}_{system}$ using (48) and (49):

$$\mathcal{D}_{system} = \left\{ (\mathbf{p}_1, \mathbf{p}_2) \in \mathbb{R}^4 \,\middle|\, \|\Delta\mathbf{p}_{12}\| \geq D_s, \mu_{12}^*, \mu_{21}^* \geq 0, \right.$$
$$\hat{\mathbf{u}}_1(\mathbf{p}_1) = \frac{1}{2}\mu_{12}^*\mathbf{a}_{12}, \hat{\mathbf{u}}_2(\mathbf{p}_2) = \frac{1}{2}\mu_{21}^*\mathbf{a}_{21},$$
$$\left. \mathbf{p}_1 \neq \mathbf{p}_{d_1}, \mathbf{p}_2 \neq \mathbf{p}_{d_2} \right\}. \quad (50)$$

where $\mathbf{a}_{12} = -\mathbf{a}_{21} = -(\mathbf{p}_1 - \mathbf{p}_2)$. From (46), we have that $\hat{\mathbf{u}}_i = \frac{1}{2}\mu_{ij}^*\mathbf{a}_{ij}$. This gives

$$\mu_{ij}^* = 2\frac{\mathbf{a}_{ij}^T\hat{\mathbf{u}}_i}{\|\mathbf{a}_{ij}\|^2} \quad (51)$$

Theorem 3. *Safety Margin Apart. In deadlock, the two robots are separated by the safety distance and the robots are on the verge of violating safety (Fig. 7(a))*

Proof. Since both robots are in deadlock, each perceives the other as active *i.e.* $\mathcal{A}_i(\mathbf{0}) = \{1, 2\}\backslash i$ for $i = \{1, 2\}$. Then, straightforward application of (45) gives this result. □

Theorem 4. $\mathcal{D}_{system}$ *is Non-Empty.* $\forall \, k_p, D_s > 0, \exists \, a$ *family of states* $(\mathbf{p}_1^*, \mathbf{p}_2^*) \in \mathcal{D}_{system}$. *These states are such that the robots and their goals are all collinear.*

Proof. To prove this theorem, we propose a set of candidate states $(\mathbf{p}_1^*, \mathbf{p}_2^*)$ and show that they satisfy the definition of deadlock (50). See Fig. 7(a) for an illustration of geometric quantities referred to in this proof.

Let $\mathbf{p}_1^* = \alpha\mathbf{p}_{d_1} + (1-\alpha)\mathbf{p}_{d_2}$ and $\mathbf{p}_2^* = \mathbf{p}_1^* - D_s\hat{e}_\beta$ where $\beta = \tan^{-1}(\frac{y_{d_2}-y_{d_1}}{x_{d_2}-x_{d_1}})$ and $\alpha \in (0, 1)$. Note that $\mathbf{p}_1^*, \mathbf{p}_2^*, \mathbf{p}_{d_1}, \mathbf{p}_{d_2}$ are collinear by construction. Then we will show that $(\mathbf{p}_1^*, \mathbf{p}_2^*) \in \mathcal{D}_{system}$. Note that

$$\mathbf{a}_{12} = -(\mathbf{p}_1^* - \mathbf{p}_2^*) = -D_s\hat{e}_\beta$$
$$\hat{\mathbf{u}}_1 = -k_{p_1}(\mathbf{p}_1^* - \mathbf{p}_{d_1}) \quad (52)$$

From definition, $\hat{e}_\beta = \frac{1}{D_G}(x_{d_2} - x_{d_1}, y_{d_2} - y_{d_1})$ where $D_G = \|\mathbf{p}_{d_2} - \mathbf{p}_{d_1}\|$ is the distance between the goals. Therefore, we have

$$\mathbf{p}_1^* - \mathbf{p}_{d_1} = -(1-\alpha)\mathbf{p}_{d_1} + (1-\alpha)\mathbf{p}_{d_2}$$
$$= (1-\alpha)D_G\hat{e}_\beta \quad (53)$$

Substituting (53) in (52) gives

$$\hat{\boldsymbol{u}}_1 = -k_{p_1}(1-\alpha)D_G \hat{e}_\beta \qquad (54)$$

From (51), (52) and (54), we deduce that μ_{12}^*

$$\mu_{12}^* = 2\frac{\boldsymbol{a}_{12}^T \hat{\boldsymbol{u}}_1}{\|\boldsymbol{a}_{12}\|_2^2} = 2k_{p_1}(1-\alpha)\frac{D_G}{D_s} > 0 \ \forall \alpha \in (0,1) \quad (55)$$

In (55) we have shown that the Lagrange multiplier μ_{12}^* is positive, which is one condition in (50). We can similarly show that $\mu_{21}^* > 0$. Further, $\hat{\boldsymbol{u}}_1 = \frac{1}{2}\mu_{12}^* \boldsymbol{a}_{12}$ is trivially satisfied because of the way we computed the multiplier in (51). Likewise, $\hat{\boldsymbol{u}}_2 = \frac{1}{2}\mu_{21}^* \boldsymbol{a}_{21}$ will be satisfied. Restricting $\alpha \in (0,1)$ ensures that $\boldsymbol{p}_i^* \neq \boldsymbol{p}_{d_i}$. Hence, the proposed states are always in deadlock as all conditions in (50) are met. $\square$

Theorem 5. $\mathcal{D}_{system}$ *is zero-measure. The system deadlock set is measure zero.*

Proof. Following the definition of $\boldsymbol{p}_1^*$ and $\boldsymbol{p}_2^*$ from theorem 3 and theorem 4, we can show that when two robots are in deadlock, their positions satisfy

$$\big\|(\boldsymbol{p}_1 - \boldsymbol{p}_{d_1})\big\| + \big\|(\boldsymbol{p}_2 - \boldsymbol{p}_{d_2})\big\| = D_s + D_G$$

This can be verified by straightforward substitution. $\square$

Characteristics of Three Robot Deadlock

Following the results shown in two robot deadlock, we now describe the three robot case. We will demonstrate that properties such as robots being on the verge of safety violation (theorem 6) and non-emptiness (theorem 7) are retained in this case as well. We are interested in analyzing *system deadlock*, which occurs when $\boldsymbol{u}_i^* = \boldsymbol{0}$ and $\hat{\boldsymbol{u}}_i \neq \boldsymbol{0}$ $\forall i \in \{1,2,3\}$. Since we are studying *system deadlock*, each robot has at-least one active constraint (each robot has two constraints in total). The *system deadlock* set $\mathcal{D}_{system}$ for three robots is defined analogously to (50).

Theorem 6. *Safety Margin Apart* *In system deadlock, either all three robots or exactly two pairs of robots are separated by the safety margin.*

Proof. The proof is kept brief because it is similar to the proof of theorem 3. Based on the number of constraints that are allowed to be active per robot, all geometric configurations can be clubbed in two categories :
Category A - This arises when all constraints of each robot are active *i.e.* $\boldsymbol{a}_{ij}^T\boldsymbol{u}_i^* = \hat{b}_{ij} = 0 \iff \|\Delta\boldsymbol{p}_{ij}\| = D_s \ \forall j \in \{1,2,3\}\backslash i \ \forall i \in \{1,2,3\}$. As a result, each robot is separated by D_s from every other robot (Fig. 8(a)).
Category B - This arises when there is exactly one robot with both its constraints active (robot i in Fig. 8(b)), and the remaining robots (j and k) have exactly one constraint active each. Hence, robot i is separated by D_s from the others. $\square$

Theorem 7. *Non-emptiness* $\forall k_p, D_s, R > 0$ *and* $\boldsymbol{p}_{d_i} = R\hat{e}_{2\pi(i-1)/3}$ *where* $i = \{1,2,3\}$, $\exists (\boldsymbol{p}_1^*, \boldsymbol{p}_2^*, \boldsymbol{p}_3^*) \in \mathcal{D}_{system}$ *where* $\boldsymbol{p}_i^*$ *are*
Category A: $\boldsymbol{p}_i^* = \frac{D_s}{\sqrt{3}}\hat{e}_{\frac{2\pi(i-1)}{3}+\pi}$ *where* $i = \{1,2,3\}$
Category B: $\boldsymbol{p}_1^* = D_s\hat{e}_\pi$, $\boldsymbol{p}_2^* = \boldsymbol{0}$, $\boldsymbol{p}_3^* = D_s\hat{e}_{\frac{\pi}{3}}$ *if robot* 2 *has both constraints active.*

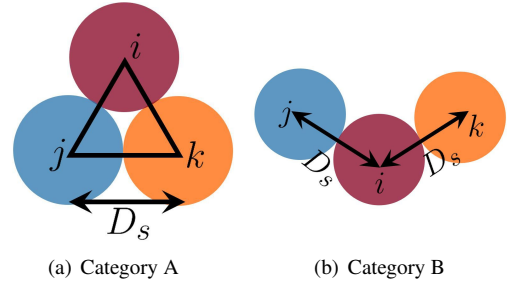(a) Category A   (b) Category B

Figure 8. Geometric configurations in three robot deadlock

Proof. This proof is similar to theorem 4 so it is skipped. $\square$

Note that in the statement of this theorem, we have predefined the desired goal positions unlike the statement of theorem 4. The candidate positions of the robots that we propose are in $\mathcal{D}_{system}$ are valid with respect to these given goals. Additionally, for category B, we proposed one set of positions that is valid in deadlock, however there is continuous family of positions that can be valid in category B which can be found in the appendix of Grover et al. [2019].

6   After Deadlock: Deadlock Resolution

We now use the properties of geometric configurations derived in section 5 to synthesize a strategy that (1) gets the robots out of deadlock, (2) ensures their safety and (3) makes them converge to their goals. One approach to achieve these objectives is to detect the incidence of deadlock while the CBF-QP controller is running on the robots and once detected, any small non-zero perturbation to the control will instantaneously give a non-zero velocity to the robots. Thereafter, CBF-QPs can take charge again and we can hope that using this controller the system state will come out of deadlock at-least for a short time. However, there is no guarantee that deadlock will not relapse because it was the CBF-QP controller that led to deadlock in the first place. Secondly, perturbations can violate safety and even lead to degraded performance. Therefore, given these limitations, we propose a controller which ensures that goal stabilization, safety and deadlock resolution are met with guarantees. We demonstrate this algorithm for the two and three robot cases. Refer to Fig. 9 for a schematic of our approach. This algorithm is described here:

1. The algorithm starts by executing controls derived from CBF-QP in Phase 1. This ensures movement of robots to the goals and safety by construction. To detect the incidence of deadlock, we continuously compare $\|\boldsymbol{u}^*\|, \|\boldsymbol{p} - \boldsymbol{p}_d\|$ against small thresholds. Once these thresholds are small enough, we declare deadlock detected and switch control to Phase 2.

2. In this phase, we rotate the robots around each other to swap positions while maintaining the safe distance. Refer to the appendix of Grover et al. [2019] for the controller derivation.

   (a) For the two-robot case, we calculate $\boldsymbol{u}_{fl}^1(t)$ and $\boldsymbol{u}_{fl}^2(t)$ using feedback linearization to ensure that $\|\Delta\boldsymbol{p}_{12}\| = D_s$ and rotation ($\dot{\theta} = -k_p(\theta -$
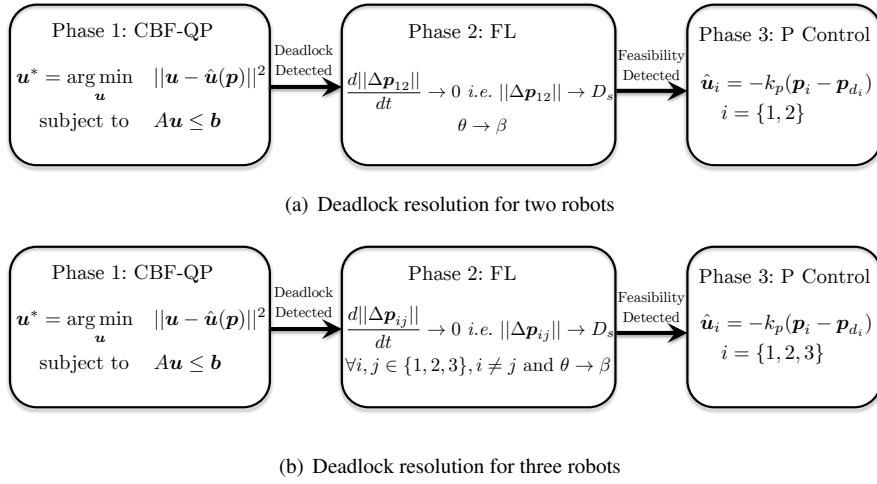
(a) Deadlock resolution for two robots

Phase 1: CBF-QP
$$\boldsymbol{u}^* = \arg\min_{\boldsymbol{u}} \quad ||\boldsymbol{u} - \hat{\boldsymbol{u}}(\boldsymbol{p})||^2$$
$$\text{subject to} \quad A\boldsymbol{u} \leq \boldsymbol{b}$$

Deadlock Detected →

Phase 2: FL
$$\frac{d||\Delta\boldsymbol{p}_{12}||}{dt} \to 0 \ i.e. \ ||\Delta\boldsymbol{p}_{12}|| \to D_s$$
$$\theta \to \beta$$

Feasibility Detected →

Phase 3: P Control
$$\hat{\boldsymbol{u}}_i = -k_p(\boldsymbol{p}_i - \boldsymbol{p}_{d_i})$$
$$i = \{1,2\}$$

(b) Deadlock resolution for three robots

Phase 1: CBF-QP
$$\boldsymbol{u}^* = \arg\min_{\boldsymbol{u}} \quad ||\boldsymbol{u} - \hat{\boldsymbol{u}}(\boldsymbol{p})||^2$$
$$\text{subject to} \quad A\boldsymbol{u} \leq \boldsymbol{b}$$

Deadlock Detected →

Phase 2: FL
$$\frac{d||\Delta\boldsymbol{p}_{ij}||}{dt} \to 0 \ i.e. \ ||\Delta\boldsymbol{p}_{ij}|| \to D_s$$
$$\forall i,j \in \{1,2,3\}, i \neq j \text{ and } \theta \to \beta$$

Feasibility Detected →

Phase 3: P Control
$$\hat{\boldsymbol{u}}_i = -k_p(\boldsymbol{p}_i - \boldsymbol{p}_{d_i})$$
$$i = \{1,2,3\}$$

Figure 9. Deadlock Resolution Algorithm Schematic

$\beta) \implies \Delta\boldsymbol{p}_{12}^T\Delta\boldsymbol{v}_{12} = 0$) (See Fig. 7(a) for θ, β). This rotation and distance invariance guarantees safety *i.e.* $h_{12} = 0$. Adding an extra constraint $\boldsymbol{u}_{fl}^1 + \boldsymbol{u}_{fl}^2 = \boldsymbol{0}$ still ensures that the problem is well posed and additionally makes the centroid static.

(b) For the three-robot case, we calculate $\boldsymbol{u}_{fl}^1(t), \boldsymbol{u}_{fl}^3(t), \boldsymbol{u}_{fl}^3(t)$ to ensure that $||\Delta\boldsymbol{p}_{12}|| = ||\Delta\boldsymbol{p}_{23}|| = ||\Delta\boldsymbol{p}_{31}|| = D_s$ and rotation $(\dot\theta = -k_p(\theta - \beta) \implies \Delta\boldsymbol{p}_{12}^T\Delta\boldsymbol{v}_{12} = 0)$ (See Fig. 7(b) for θ, β). This guarantees safety *i.e.* $h_{12} = h_{23} = h_{31} = 0$. Similarly, we impose $\boldsymbol{u}_{fl}^1 + \boldsymbol{u}_{fl}^2 + \boldsymbol{u}_{fl}^3 = \boldsymbol{0}$ to make the centroid static. Note that both in simulations and experiments, we observed incidence of only Category A deadlock so our resolution algorithm is specific for this category.

3. Once the robots swap their positions, their new positions will ensure that prescribed proportional controllers will be feasible in the future. Thus, after convergence of Phase 2, control switches to Phase 3, which simply uses the prescribed controllers. This phase guarantees that the distance between robots is non-decreasing and safety is maintained as we prove in theorem 8.

Fig. 10 shows simulation results from running this strategy on two robots. We also conducted an experimental validation of this algorithm using Khepera 4 nonholonomic robots. The videos for experimental results can be found at https://bit.ly/3hBJIew. Note that for nonholnomic robots, we noticed from experiments and simulations (in two robot case) that deadlock only occurs if the body frames of both robots are perfectly aligned with one another at $t = 0$. Since this alignment is difficult to establish in experiments due to sensor noise, we simulated a virtual deadlock at $t = 0$ *i.e.* assumed that the initial position of robots are in deadlock.

Figs. 11 show the simulation results from running this strategy on three robots. Fig. 12 shows snapshots of this video during different phases of the execution of the deadlock resolution algorithm. In three robots, we **did** observe deadlock experimentally as can be seen here https:

(a) Robot 1 Position   (b) Robot 2 Position

Figure 10. Positions of robots from deadlock resolution algorithm for two robots. The vertical bars separate the three phases. In all figures, final positions converge to desired positions in Phase 3. Simulation video at https://youtu.be/nrWXdn_3nI4 and experimental at https://bit.ly/3hBJIew
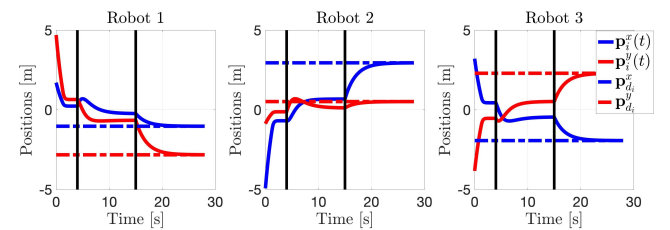
Figure 11. Positions of robots from deadlock resolution algorithm for three robots. The vertical bars separate the three phases. Final positions converge to desired positions in Phase 3. Simulation video at https://youtu.be/5xbx4mk27xc and experimental validation at https://youtu.be/IqYmkQUjrBI

//youtu.be/e6eOeuh7Uec. We only observed incidence of Category A deadlock in both simulations and experiments so our resolution algorithm is designed for this category. The videos for experimental results for validation of the resolution algorithm can be found at https://youtu.be/IqYmkQUjrBI. The video demonstrates that our resolution algorithm was able to safely deliver the robots to their goals. We next prove that this strategy ensures resolution
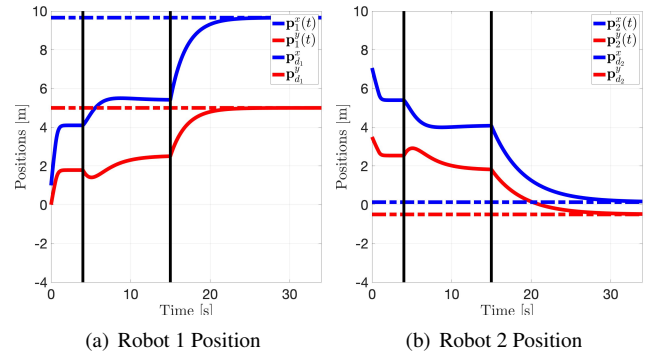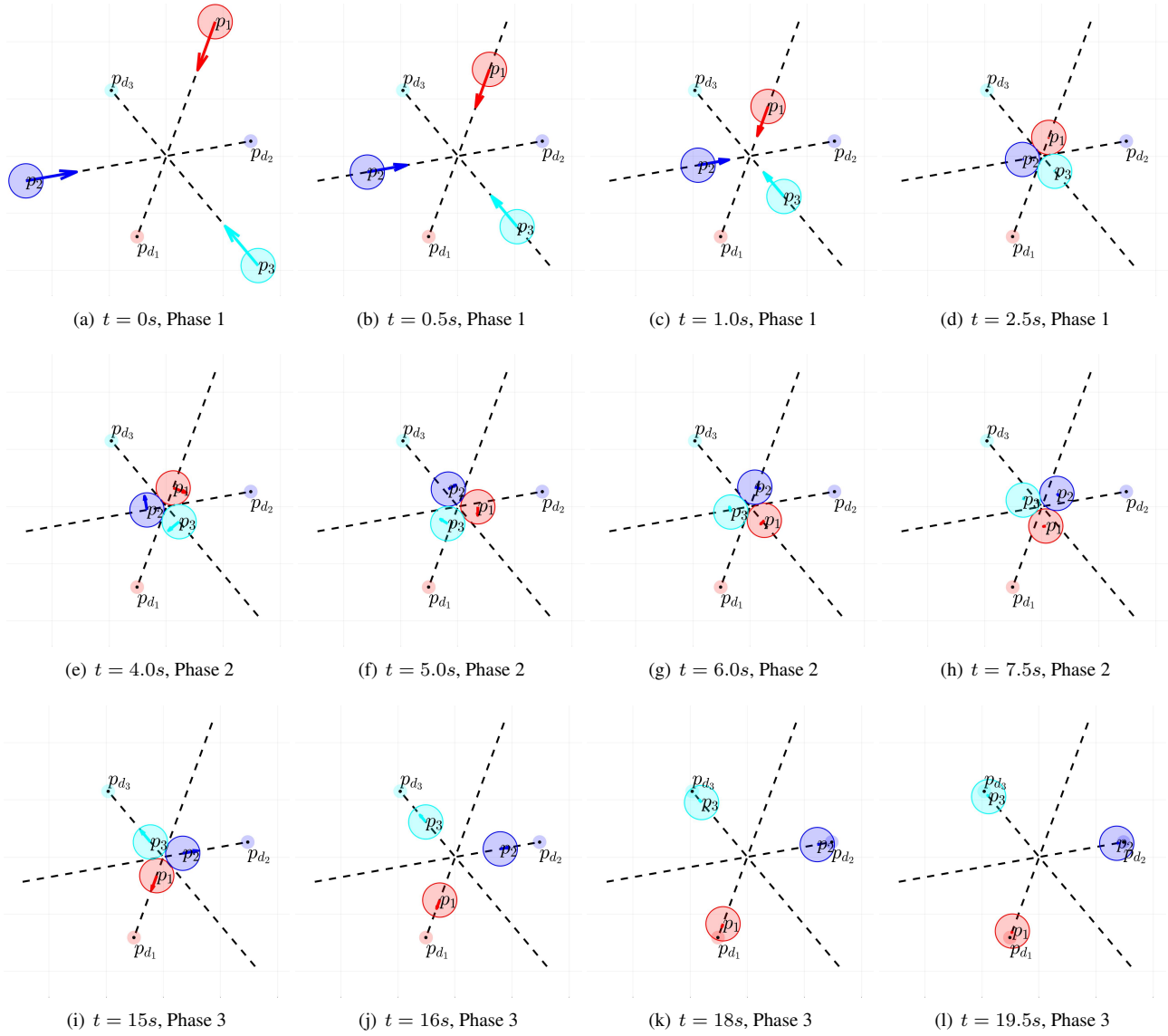
(a) $t = 0s$, Phase 1   (b) $t = 0.5s$, Phase 1   (c) $t = 1.0s$, Phase 1   (d) $t = 2.5s$, Phase 1

(e) $t = 4.0s$, Phase 2   (f) $t = 5.0s$, Phase 2   (g) $t = 6.0s$, Phase 2   (h) $t = 7.5s$, Phase 2

(i) $t = 15s$, Phase 3   (j) $t = 16s$, Phase 3   (k) $t = 18s$, Phase 3   (l) $t = 19.5s$, Phase 3

Figure 12. Snapshots from running the deadlock resolution algorithm on three robots. The top row corresponds to Phase 1 where we use CBF-QPs. The middle row corresponds to Phase 2 that uses feedback linearization for rotating the robots and the last row corresponds to Phase 3 that uses proportional controls. The simulation video can be watched at https://youtu.be/5xbx4mk27xc

of deadlock and convergence of robots to their goals by exploiting the properties we derived in theorem 3 and theorem 4. We prove this theorem for $N = 2$ since $N = 3$ is a trivial extension of $N = 2$.

Theorem 8. *The three phase strategy ensures that the robots will not fall back in deadlock and will converge to their goals.*

Proof. See appendix 9.

7 Conclusions

In this paper, we analyzed the characteristic properties of deadlock that results from using CBF based QPs for avoidance control in multirobot systems. We broke our analysis into three phases consistent with the chronology of deadlock incidence. For before deadlock, we demonstrated that symmetry in initial and goal positions leads to deadlocks and heterogeneity in controller parameters may not be enough to evade deadlock. The key lesson that we

learned from the examples we considered is that symmetry will always lead to deadlock resulting in robots grazing each other asymptotically. Thus a symmetry breaking controller with some random noise can potentially prevent such configurations from being attained. However even with randomness, deadlock avoidance cannot be taken for granted. Thus, a guaranteed way to resolve deadlock is to allow it to occur and then correct it using a reactive mechanism. However, devising this corrective strategy necessitates knowing the geometric characteristics of the robot configurations when the system is in deadlock.

In that spirit, we used the KKT conditions of the CBF-QPs and analyzed the geometric properties of an N robot arrangement when it has fallen in deadlock. We demonstrated how to interpret deadlock as a subset of the state space and proved that in deadlock, the robots are on the verge of violating safety. Additionally, we showed that this set is non-empty and bounded. We also demonstrated that the number of valid geometric configurations in deadlock

increases approximately exponentially with the number of robots which makes the analysis and resolution for $N \geq 4$ complex. Finally, by using these properties, we devised a corrective control algorithm to force the robots out of deadlock and ensure task completion.

There are several directions we would like to explore in future. Firstly, we want to extend this to $N \geq 4$ case. In the $N = 4$ case, we determined a large number of admissible geometric configurations. We found that there exist bijections among some of these configurations depending on the number of total active constraints. We believe this property can be exploited to reduce the complexity down to the equivalence classes of these bijections. We will exploit this line of approach to simplify analysis for $N \geq 4$ cases. Secondly, while in this work we focused on CBF-QPs for analysis, our developed theory can be used to analyze properties of deadlock resulting from other optimization based such as velocity obstacles and the safe-set method. So in future, we will evaluate the generalizabilty of our derived results specifically the geometric properties of robot arrangements across the spectrum of reactive optimization based controllers, and explore properties that make a particular algorithm immune to deadlock.

A   Appendix

Two Robot Deadlock Results

Lemma A.1. *If at $t = 0$, $D_{init} > \beta_+^i$, then $\boldsymbol{u}_i^*(0) = \hat{\boldsymbol{u}}_i(0)$, i.e. the collision avoidance constraint of robot i is inactive.*

Proof. To determine the control returned by the optimization for robot one at $t = 0$, we compute the value of the flag $f_{12}(\hat{\boldsymbol{u}}_1(0))$. This flag evaluates whether $\hat{\boldsymbol{u}}_1(0)$ is a feasible controller *i.e.* it prevents collision of robot 1 with 2. Based on the initial positions and goals from (17)-(18), we have

$$
\begin{aligned}
f_{12}(\hat{\boldsymbol{u}}_1(0)) &= \boldsymbol{a}_{12}^T(0)\hat{\boldsymbol{u}}_1(0) - b_{12}(0) \\
&= k_{p_1}(\boldsymbol{p}_1(0) - \boldsymbol{p}_2(0))^T(\boldsymbol{p}_1(0) - \boldsymbol{p}_{d_1}) \\
&\quad - \frac{\gamma}{4}(D_{init.}^2 - D_s^2) \\
&= -\frac{\gamma}{4}D_{init}^2 + D_{G_1}k_{p_1}D_{init} + \frac{\gamma}{4}D_s^2 \\
&\coloneqq g_{12}(D_{init.})
\end{aligned}
\tag{56}
$$

Define $g_{12}(D_{init.}) \coloneqq f_{12}(\hat{\boldsymbol{u}}_1(0))$ to emphasize dependence on $D_{init.}$, the initial distance between the robots. (56) is a quadratic polynomial in D_{init} and has two zeros at

$$
\beta_\pm^1 = \frac{2D_{G_1}k_{p_1}}{\gamma} \pm \sqrt{\left(\frac{2D_{G_1}k_{p_1}}{\gamma}\right)^2 + D_s^2}
\tag{57}
$$

where the subscript of β indicates the sign of β. Now, since the graph of $g_{12}(D_{init})$ is a downward facing parabola, we know that $g_{12}(D_{init}) < 0 \ \forall \ D_{init} \in (-\infty, \beta_-^1) \cup (\beta_+^1, \infty) \implies g_{12}(D_{init}) < 0 \ \forall \ D_{init} \in (\beta_+^1, \infty)$. We call β_+^1 to be the critical distance for robot 1. If at $t = 0$, the distance between the robots is such that $D_{init} > \beta_+^1$, then at $t = 0$, $g_{12}(D_{init.}) < 0 \iff f_{12}(\hat{\boldsymbol{u}}_1(0)) < 0 \implies \boldsymbol{u}_1^*(0) = \hat{\boldsymbol{u}}_1(0)$. Similarly, we can compute $g_{21}(D_{init.}) \coloneqq f_{21}(\hat{\boldsymbol{u}}_2(0))$ which has roots at $\beta_\pm^2$. Hence, if at $t = 0$, $D_{init} > \beta_+^2$, then $\boldsymbol{u}_2^*(0) = \hat{\boldsymbol{u}}_2(0)$. $\square$

Lemma A.2. *$\exists$ a finite time t_1 as described in Def. (24), until which the collision avoidance constraints of both robots are simultaneously inactive.*

Proof. Since at $t = 0$, both robots use their prescribed controls $\hat{\boldsymbol{u}}_i$, let us assume that they continue to do so for the interim. Therefore, the dynamics of the robots are

$$
\dot{\boldsymbol{p}}_i = \hat{\boldsymbol{u}}_i = -k_{p_i}(\boldsymbol{p}_i - \boldsymbol{p}_{d_i}).
\tag{58}
$$

(58) can be integrated analytically using the initial and desired positions of the two robots from (17) and (18) to give

$$
\begin{aligned}
\boldsymbol{p}_1(t) &= \boldsymbol{p}_{d_1} - D_{G_1}e^{-k_{p_1}t}\hat{\boldsymbol{e}}_\alpha \\
\boldsymbol{p}_2(t) &= \boldsymbol{p}_{d_2} + D_{G_2}e^{-k_{p_2}t}\hat{\boldsymbol{e}}_\alpha
\end{aligned}
\tag{59}
$$

We can compute the relative position between 1 and 2 as:

$$
\begin{aligned}
\Delta\boldsymbol{p}_{21}(t) &= (D_{G_1}e^{-k_{p_1}t} + D_{G_2}e^{-k_{p_2}t} - K)\hat{\boldsymbol{e}}_\alpha \\
&= D(t)\hat{\boldsymbol{e}}_\alpha
\end{aligned}
\tag{60}
$$

where $K \coloneqq D_{G_1} + D_{G_2} - D_{init}$. $D(t)$ denotes the inter-robot distance as a function of time for the duration in which both robots continue to use their proportional controllers. From (60), it is given by

$$
D(t) \coloneqq (D_{G_1}e^{-k_{p_1}t} + D_{G_2}e^{-k_{p_2}t}) - K,
\tag{61}
$$

The nominal controls can are given by

$$
\hat{\boldsymbol{u}}_1(t) = -k_{p_1}(\boldsymbol{p}_1(t) - \boldsymbol{p}_{d_1}) \quad = +k_{p_1}D_{G_1}e^{-k_{p_1}t}\hat{\boldsymbol{e}}_\alpha \tag{62}
$$
$$
\hat{\boldsymbol{u}}_2(t) = -k_{p_2}(\boldsymbol{p}_2(t) - \boldsymbol{p}_{d_2}) \quad = -k_{p_2}D_{G_2}e^{-k_{p_2}t}\hat{\boldsymbol{e}}_\alpha \tag{63}
$$

Now, we use (61)-(63) to compute the feasibility flag for robot 1 using (22)

$$
\begin{aligned}
f_{12}(\hat{\boldsymbol{u}}_1(t)) &= \boldsymbol{a}_{12}^T(t)\hat{\boldsymbol{u}}_1(t) - b_{12}(t) \\
&= -\frac{\gamma}{4}D^2(t) + k_{p_1}D_{G_1}e^{-k_{p_1}t}D(t) + \frac{\gamma}{4}D_s^2
\end{aligned}
\tag{64}
$$

which is a downwards facing parabola in $D(t)$. Thus, $f_{12}(\hat{\boldsymbol{u}}_1(t)) < 0$ for $D(t) > \beta_+^1(t)$ where

$$
\beta_+^1(t) = \frac{2D_{G_1}k_{p_1}e^{-k_{p_1}t}}{\gamma} + \sqrt{\left(\frac{2D_{G_1}k_{p_1}e^{-k_{p_1}t}}{\gamma}\right)^2 + D_s^2}
\tag{65}
$$

and likewise we can define $\beta_+^2(t)$ for robot two by replacing k_{p_1}, D_{G_1} with k_{p_2}, D_{G_2} in (65). Note that $D(t)$, $\beta_+^1(t)$ and $\beta_+^2(t)$ are monotonically decreasing with time. Additionally, recall from (17) that $D(0) = D_{init}$ and from Assumption 1 that $D_{init.} > \beta_+^1$. However, while $D(t)$ converges to $-(D_{G_1} + D_{G_2} - D_{init}) < 0$, $\beta_+^1(t), \beta_+^2(t)$ converge to $D_s > 0$. Therefore, there exists a time t_a at which $D(t_a) = \beta_+^1(t_a) \implies f_{12}(\hat{\boldsymbol{u}}_1(t_a)) = 0$ and a time t_b at which $D(t_b) = \beta_+^2(t_b) \implies f_{21}(\hat{\boldsymbol{u}}_2(t_b)) = 0$. We assume WLOG that $\beta_+^1(t_a) > \beta_+^2(t_a) \implies t_a < t_b$ and hence we define:

$$
t_1 \coloneqq \min\{t_a, t_b\} = t_a
\tag{66}
$$

t_a is the time at which the collision avoidance constraint of robot one becomes active *i.e.* $\boldsymbol{a}_{12}^T(t_a)\boldsymbol{u}_1^*(t_a) = b_{12}(t_a)$

while the collision avoidance constraint of robot two is still inactive *i.e.* $\boldsymbol{a}_{21}^T(t_a)\hat{\boldsymbol{u}}_2(t_a) < b_{21}(t_a)$. Thus t_a is the time in Def. 24, *i.e.* the maximum time until which both robots' constraints are inactive. See Fig. 4 for an illustration of phase 1 which ends at $t = t_1$ when $D(t) = \beta_+^1(t)$. Finally, $D(t)$ is monotonically decreasing $\forall t$ in (60) and $D(0) = D_{init} > 0$, $D(t_1) = \beta_+^1(t_1) > 0$, therefore, $D(t) > 0 \ \forall t \in [0, t_1)$. From (60), note that the distance between the robots is $\|\Delta \boldsymbol{p}_{21}(t)\| := |D(t)| = D(t)$ $\forall t \in [0, t_1)$. Therefore, $\|\Delta \boldsymbol{p}_{21}(t)\|$ is also monotonically *decreasing* until t_1. □

Lemma A.3. $\exists$ *a finite time t_2 as in Def.* (27), *until which the constraint of robot two stays inactive.*

Proof. Let us assume that there is a time t_2 until which the collision avoidance constraint of robot two remains inactive. Therefore, the dynamics of robot two are governed by the prescribed nominal control. Thus,

$$\dot{\boldsymbol{p}}_2 = \hat{\boldsymbol{u}}_2 = -k_{p_2}(\boldsymbol{p}_2 - \boldsymbol{p}_{d_2})$$
$$\implies \boldsymbol{p}_2(t) = \boldsymbol{p}_{d_2} + D_{G_2}e^{-k_{p_2}t}\hat{\boldsymbol{e}}_\alpha \qquad (67)$$

On the other hand, the control input to robot one is

$$\boldsymbol{u}_1^* = \hat{\boldsymbol{u}}_1 - \frac{1}{2}\mu_{12}\boldsymbol{a}_{12}, \qquad (68)$$

where $\mu_{12} \neq 0$ because the collision avoidance constraint of robot one became active at t_1 *i.e.* $\boldsymbol{a}_{12}^T(t)\boldsymbol{u}_1^*(t) = b_{12}(t) \ \forall t \geq t_1$. It's expression is given as follows

$$\mu_{12} = 2\frac{\boldsymbol{a}_{12}^T\hat{\boldsymbol{u}}_1 - b_{12}}{\|\boldsymbol{a}_{12}\|^2} \qquad (69)$$

Therefore, the dynamics of robot one $\forall t \geq t_1$ are

$$\dot{\boldsymbol{p}}_1 = \hat{\boldsymbol{u}}_1 - \frac{\boldsymbol{a}_{12}^T\hat{\boldsymbol{u}}_1 - b_{12}}{\|\boldsymbol{a}_{12}\|^2}\boldsymbol{a}_{12}$$
$$= \underbrace{\hat{\boldsymbol{u}}_1 - \frac{\boldsymbol{a}_{12}^T\hat{\boldsymbol{u}}_1}{\|\boldsymbol{a}_{12}\|^2}\boldsymbol{a}_{12}}_{\boldsymbol{u}_\perp} + \underbrace{\frac{b_{12}}{\|\boldsymbol{a}_{12}\|^2}\boldsymbol{a}_{12}}_{\boldsymbol{u}_\|} \qquad (70)$$

On closer inspection, note that $\boldsymbol{u}_\perp \perp \boldsymbol{a}_{12} \iff \boldsymbol{u}_\perp \perp -\Delta \boldsymbol{p}_{12}$. We want to derive an analytical expression of $\boldsymbol{p}_1(t)$ just like we did for robot 2 in (67). However, it is difficult to analytically integrate (70). Instead, we will show via recursion that the $\boldsymbol{p}_1(t)$ can be expressed as

$$\boldsymbol{p}_1(t) = \boldsymbol{p}_{d_1} - D_{G_1}\eta(t)\hat{\boldsymbol{e}}_\alpha \qquad (71)$$

for some function $\eta(t) \ \forall t \geq t_1$. This expression is valid at $t = t_1$ for $\eta(t_1) = e^{-k_{p_1}t_1}$ as shown in (59). This representation highlights that the position of robot one is confined along $\hat{\boldsymbol{e}}_\alpha$. We will show that this property is maintained throughout phase 2 because the component of velocity input to robot one perpendicular to $\hat{\boldsymbol{e}}_\alpha$ will vanish (which will become the cause of *deadlock*). Recall at $t = t_1$

$$\boldsymbol{a}_{12}(t_1) = \Delta \boldsymbol{p}_{21}(t_1) = D(t_1)\hat{\boldsymbol{e}}_\alpha \qquad (72)$$

Therefore, $\boldsymbol{u}_\perp(t_1) = \hat{\boldsymbol{u}}_1 - \dfrac{\boldsymbol{a}_{12}^T\hat{\boldsymbol{u}}_1}{\|\boldsymbol{a}_{12}\|^2}\boldsymbol{a}_{12}$

$$= k_{p_1}D_{G_1}\eta(t_1)\hat{\boldsymbol{e}}_\alpha$$
$$\quad - \frac{D(t_1)\hat{\boldsymbol{e}}_\alpha^T(k_{p_1}D_{G_1}\eta(t_1)\hat{\boldsymbol{e}}_\alpha)}{D^2(t_1)}(D(t_1)\hat{\boldsymbol{e}}_\alpha)$$
$$= \boldsymbol{0}$$
$$\boldsymbol{u}_\|(t_1) = \frac{b_{12}}{\|\boldsymbol{a}_{12}\|^2}\boldsymbol{a}_{12}$$
$$= \gamma\frac{D^2(t_1) - D_s^2}{4D(t_1)}\hat{\boldsymbol{e}}_\alpha \qquad (73)$$

Thus, integrating the velocity for a small time step gives

$$\boldsymbol{p}_1(t_1 + \Delta t) = \boldsymbol{p}_1(t_1) + \Delta t\big(\boldsymbol{u}_\perp(t_1) + \boldsymbol{u}_\|(t_1)\big)$$
$$= \boldsymbol{p}_{d_1} - D_{G_1}\left(\eta(t_1) - \Delta t\gamma\frac{D^2(t_1) - D_s^2}{4D_{G_1}D(t_1)}\right)\hat{\boldsymbol{e}}_\alpha$$
$$= \boldsymbol{p}_{d_1} - D_{G_1}\eta(t_1 + \Delta t)\hat{\boldsymbol{e}}_\alpha \qquad (74)$$

Through (74), we have demonstrated that the updated position of robot one admits the general form given by (71) because the perpendicular component of the velocity vanishes. As a result, the robot never acquires any displacement along the perpendicular component. Hence, the dynamics of robot one are always

$$\dot{\boldsymbol{p}}_1 = \boldsymbol{u}_\| = \gamma\frac{\|\Delta \boldsymbol{p}_{21}\|^2 - D_s^2}{4\|\Delta \boldsymbol{p}_{21}\|^2}\Delta \boldsymbol{p}_{21}, \qquad (75)$$

$\forall t \geq t_1$. The relative dynamics are

$$\dot{\Delta \boldsymbol{p}}_{21} = -\gamma\frac{\|\Delta \boldsymbol{p}_{21}\|^2 - D_s^2}{4\|\Delta \boldsymbol{p}_{21}\|^2}\Delta \boldsymbol{p}_{21} - k_{p_2}(\boldsymbol{p}_2 - \boldsymbol{p}_{d_2}) \qquad (76)$$

Let $\Delta \boldsymbol{p}_{21}(t) = D(t)\hat{\boldsymbol{e}}_\alpha$, then we get for $t \geq t_1$

$$\dot{D}(t) = -\gamma\frac{D^2(t) - D_s^2}{4D(t)} - k_{p_2}D_{G_2}e^{-k_{p_2}t} \qquad (77)$$

where $D(t_1) = \beta_+^1(t_1) > D_s$. Note that for $D(t) > D_s$, $\dot{D}(t) < 0 \implies D(t)$ is monotonically decreasing at least until t_c where $t_c := \{t | D(t) = D_s\}$. Recall from phase 1 that the collision avoidance constraint of robot two is inactive *i.e.* $f_{21}(\hat{\boldsymbol{u}}_2(t)) < 0$ as long as $D(t) > \beta_+^2(t)$. Additionally, recall that $\beta_+^2(t)$ is monotonically decreasing with respect to time and converges to D_s. Moreover, from phase 1, recall that $D(t_1) = \beta_+^1(t_1) > \beta_+^2(t_1)$. Hence, there exists a time $t_2 \leq t_c$ at which $D(t) = \beta_+^2(t)$. This time t_2 is precisely the time in (27). See Fig. 4 for an illustration of phase 2 which ends at $t = t_2$ when $D(t) = \beta_+^2(t)$. This concludes phase 2 and marks the start of phase 3. □

Lemma A.4. *The distance between robots converges to the safety margin D_s at which point they stop moving and fall in deadlock.*

Proof. Letting $\Delta \boldsymbol{p}_{21}(t) = D(t)\hat{\boldsymbol{e}}_\alpha$ in (31)

$$\dot{D}(t) = -\gamma\frac{D^2(t) - D_s^2}{2D(t)} \qquad (78)$$

where $D(t_2) = \beta_+^2(t_2) > D_s$. The solution to this is

$$D(t) = \sqrt{(D^2(t_2) - D_s^2)e^{-\gamma(t-t_2)} + D_s^2} \quad (79)$$

Therefore $D(t) \longrightarrow D_s$ and so $\|\Delta\boldsymbol{p}_{12}(t)\| \longrightarrow D_s$. Therefore, the robots will be just at the verge of colliding. Moreover, note from (29) that as $\|\Delta\boldsymbol{p}_{12}(t)\| \longrightarrow D_s$, $\dot{\boldsymbol{p}}_1 \longrightarrow \boldsymbol{0}$, $\dot{\boldsymbol{p}}_2 \longrightarrow \boldsymbol{0}$, *i.e.* the robots stop moving. However, once they stop, note that $\lim_{t \longrightarrow \infty} \Delta\boldsymbol{p}_{21}(t) = D_s\hat{\boldsymbol{e}}_\alpha$, yet $\Delta\boldsymbol{p}_{d_{21}} = -(D_{G_1} + D_{G_2} - D_{init.})\hat{\boldsymbol{e}}_\alpha$. Therefore, the robots are not at their goals (because the goal vector is anti-parallel to the vector connecting the robots). Hence, we conclude that the robots have fallen in *deadlock*. $\square$

Three Robot Deadlock Results

Lemma A.5. *If $D_{init.} > \beta_+$, then $\boldsymbol{u}_i^*(0) = \hat{\boldsymbol{u}}_i(0) \ \forall i$*

Proof. To keep calculations brief, we compute the flag $f_{ij}(\hat{\boldsymbol{u}}_i(0))$ for $i = 1, j = 2$. Using (32)-(33), we have

$$f_{12}(\hat{\boldsymbol{u}}_1(0))) = \boldsymbol{a}_{12}^T(0)\hat{\boldsymbol{u}}_1(0) - b_{12}(0)$$
$$= -\frac{\gamma}{4}D_{init}^2 + \frac{\sqrt{3}D_G k_p}{2}D_{init} + \frac{\gamma}{4}D_s^2$$
$$:= g(D_{init.}) \quad (80)$$

We can show that $f_{ij}(\hat{\boldsymbol{u}}_i) = g(D_{init.}) \ \forall i \in \{1,2,3\}, j \neq i$ because of the equilateral configuration at $t = 0$. $g(D_{init.})$ is a quadratic polynomial in $D_{init.}$ with a zero at β_+:

$$\beta_+ = \frac{\sqrt{3}D_G k_p}{\gamma} + \sqrt{\left(\frac{\sqrt{3}D_G k_p}{\gamma}\right)^2 + D_s^2} \quad (81)$$

Therefore, $\qquad D_{init} > \beta_+ \implies g(D_{init.}) < 0 \qquad \Longleftrightarrow$
$f_{ij}(\hat{\boldsymbol{u}}_i(0)) < 0 \ \forall i \in \{1,2,3\}, j \in \{1,2,3\}\backslash i, \qquad \Longleftrightarrow$
$\boldsymbol{u}_i^*(0) = \hat{\boldsymbol{u}}_i(0) \ \forall i \in \{1,2,3\}$. $\square$

Lemma A.6. *All robots continue to move in an equilateral triangular configuration with centroid fixed at $\boldsymbol{c}(0)$*

Proof. Using (36), the centroid's position is given by

$$\boldsymbol{c}(t) = \underbrace{\frac{1}{3}\sum_{i=1}^3(\boldsymbol{p}_i(0) - \boldsymbol{p}_{d_i})e^{-k_p t}}_{\boldsymbol{0}} + \underbrace{\frac{1}{3}\sum_{i=1}^3\boldsymbol{p}_{d_i}}_{\boldsymbol{c}(0)}$$
$$= \boldsymbol{c}(0) \quad (82)$$

Here the two terms are zero because of the way we defined the initial position and goal locations in (32)-(33). Next, to show that the three robots move along the vertices of an equilateral triangle, we compute the distance between robots i and j and show that it identical for every pair. From (36),

$$\|\Delta\boldsymbol{p}_{ij}(t)\| = \sqrt{\underbrace{\|\Delta\boldsymbol{p}_{ij}(0)\|^2}_{\text{term 1}}e^{-2k_p t} + \underbrace{\|\Delta\boldsymbol{p}_{d_{ij}}\|^2}_{\text{term 2}}(1-e^{-k_p t})^2}$$
$$+ 2\underbrace{\Delta\boldsymbol{p}_{ij}^T(0)\Delta\boldsymbol{p}_{d_{ij}}}_{\text{term 3}}(e^{-k_p t} - e^{-2k_p t}) \quad (83)$$

One can show that term 1, term 2 and term 3 are identical for all $i, j \in \{1,2,3\}, j \neq i$ using (32), (33). Therefore, the

distance of robot i from robot $j \neq i$ is same for all $i \in \{1,2,3\}$, hence the robots move along the vertices of an equilateral triangle. A second invariant is the angle made by the vector connecting robots i, j with the X_w axis of the world. This can be shown by demonstrating that $\Delta\boldsymbol{p}_{ij}(t)$ remains parallel to $\Delta\boldsymbol{p}_{ij}(0)$ as follows

$$\Delta\boldsymbol{p}_{ij}(t) \times \Delta\boldsymbol{p}_{ij}(0) = \underbrace{(\Delta\boldsymbol{p}_{ij}(0) \times \Delta\boldsymbol{p}_{ij}(0))}_{0}e^{-k_p t}$$
$$+ \underbrace{(\Delta\boldsymbol{p}_{d_{ij}} \times \Delta\boldsymbol{p}_{ij}(0))}_{\text{Term 1}}(1 - e^{-k_p t})$$
$$= \boldsymbol{0} \quad (84)$$

Term 1 vanishes because $\Delta\boldsymbol{p}_{d_{ij}}$ is anti-parallel to $\Delta\boldsymbol{p}_{ij}(0)$ using (32), (33). Since now we have shown that the three robots move along the vertices of an equilateral triangle, the positions of the robots can still be written in the form similar to the way we defined their initial positions in (32):

$$\boldsymbol{p}_1(t) = e^{-k_p t}\boldsymbol{p}_1(0) + (1 - e^{-k_p t})\boldsymbol{p}_{d_1}$$
$$\boldsymbol{p}_2(t) = \boldsymbol{p}_1(t) + \tilde{D}(t)\hat{\boldsymbol{e}}_\alpha$$
$$\boldsymbol{p}_3(t) = \boldsymbol{p}_2(t) + \tilde{D}(t)\hat{\boldsymbol{e}}_{\alpha+\frac{2\pi}{3}}, \quad (85)$$

where $\tilde{D}(0) = D_{init.}$. $\tilde{D}(t)$ denotes the inter-robot distance between any pair of robots. Here, the angle between $\Delta\boldsymbol{p}_{21}$ and X_w is still α as in (32) because of (84). Thus, the robots move along the vertices of an equliateral triangle using $\hat{\boldsymbol{u}}_i(t) \ \forall i \in \{1,2,3\}$. *This symmetry is because (1) $\hat{\boldsymbol{u}}_i(t)$ has identical gains (i.e. k_p) $\forall i$ and (2) the distance of initial position to goal is identical for all robots (D_G).* $\square$

Note that in (85), $\tilde{D}(t)$ is

$$\tilde{D}(t) = \hat{\boldsymbol{e}}_\alpha^T \Delta\boldsymbol{p}_{21}(t)$$
$$= \hat{\boldsymbol{e}}_\alpha^T\left(\Delta\boldsymbol{p}_{21}(0)e^{-k_p t} + \Delta\boldsymbol{p}_{d_{21}}(1 - e^{-k_p t})\right)$$
$$= (D_{init.} - \sqrt{3}D_G) + \sqrt{3}D_G e^{-k_p t} \quad (86)$$

Using this definition of $\tilde{D}(t)$, we now demonstrate that there exists a time when all the collision avoidance constraints will inevitably become active.

Lemma A.7. *There exists a time t_{ij} when $f_{ij}(\hat{\boldsymbol{u}}_i(t)) = \boldsymbol{a}_{ij}^T(t)\hat{\boldsymbol{u}}_i(t) - b_{ij}(t) = 0$ i.e. when the collision avoidance constraint of robot i with robot j becomes active. Furthermore, t_{ij} is identical $\forall i \in \{1,2,3\}, j \in \{1,2,3\}\backslash i$*

Proof. Using (36), we can find that $f_{ij}(\hat{\boldsymbol{u}}_i(t))$ (as a function of time) is identical for $\forall i \in \{1,2,3\}, j \in \{1,2,3\}\backslash i$. This is again due to the symmetry in the positions at time t. Now, we evaluate $f_{ij}(\hat{\boldsymbol{u}}_i)$ as a function of $\tilde{D}$. For brevity, we evaluate this for $i = 1, j = 2$:

$$g(\tilde{D}) := f_{12}(\hat{\boldsymbol{u}}_1) = \boldsymbol{a}_{12}^T\hat{\boldsymbol{u}}_1 - b_{12}$$
$$= -\frac{\gamma}{4}\tilde{D}^2(t) + \frac{\sqrt{3}}{2}k_p D_G e^{-k_p t}\tilde{D}(t) + \frac{\gamma}{4}D_s^2 \quad (87)$$

Note that $g(\tilde{D})$ is quadratic in $\tilde{D}$ with a zero at

$$\beta_+(t) = \frac{\sqrt{3}D_G k_p e^{-k_p t}}{\gamma} + \sqrt{\left(\frac{\sqrt{3}D_G k_p e^{-k_p t}}{\gamma}\right)^2 + D_s^2} \quad (88)$$

Therefore, if $\tilde{D}(t) \leq \beta_+(t) \implies f_{12}(\hat{\boldsymbol{u}}_1(t)) \geq 0$. Now, note from (86) that $\tilde{D}(t)$ is monotonically decreasing with t, $\tilde{D}(0) = D_{init.} > \beta_+$ (from Assumption (3)). Also, note that $\lim_{t \to \infty} \tilde{D}(t) = D_{init.} - \sqrt{3}D_G < 0$ from Assumption (2). Similarly, $\beta_+(t)$ is monotonically decreasing. $\beta_+(0) < D_{init.}$ and $\lim_{t \to \infty} \beta_+(t) = D_s > 0$. Therefore, there exists a time when $\tilde{D}(t)$ intersects $\beta_+(t)$ i.e. $t_1 = \{t | \tilde{D}(t) = \beta_+(t)\}$. This is equivalent to $t_1 = \{t | f_{ij}(\hat{\boldsymbol{u}}_i(t)) = 0 \; \forall i \in \{1, 2, 3\}, \; j \in \{1, 2, 3\} \backslash i\}$. This is precisely when constraints of **all** robots become active. $\quad\square$

Lemma A.8. *The distance between the robots converges to the safety margin D_s and the robots stop moving and fall in deadlock.*

Proof. Noting from (38) that $\Delta \boldsymbol{p}_{12}(t) = -D(t)\hat{\boldsymbol{e}}_\alpha$ and combining with (40), we deduce that

$$\dot{D}(t) = -\gamma \frac{D^2(t) - D_s^2}{2D(t)} \qquad \forall t \geq t_1$$

$$\implies D(t) = \sqrt{(D^2(t_1) - D_s^2)e^{-\gamma(t-t_1)} + D_s^2}$$

$$\implies \lim_{t \to \infty} D(t) = D_s \tag{89}$$

Hence, it follows from (39) that $\dot{\boldsymbol{p}}_i \longrightarrow \boldsymbol{0} \; \forall i \in \{1, 2, 3\}$. Moreover, note that $\lim_{t \to \infty} \Delta \boldsymbol{p}_{12}(t) = -D_s \hat{\boldsymbol{e}}_\alpha$ yet $\Delta \boldsymbol{p}_{d_{12}} = (\sqrt{3}D_G - D_{init.})\hat{\boldsymbol{e}}_\alpha \neq -D_s \hat{\boldsymbol{e}}_\alpha$ (from (33) and Assumption 2). Therefore, the robots are not at their goals, and static, thus they have fallen in *deadlock*. $\quad\square$

Deadlock Resolution Algorithm Proof

Theorem 9. *The three phase strategy ensures that the robots will not fall back in deadlock and will converge to their goals.*

Proof. We would like to show that once phase three control begins, the robots will never fall back in deadlock. We will do this by showing that the distance between the robots is non-decreasing, once phase three control starts. We make one assumption which is needed for a technicality in the proof but is easily achieved in practice.

Assumption 4. $D_G > D_s$ *i.e. the inter-goal distance is greater than the safety margin.*

This is required because otherwise, the robots will never be at-least D_s apart while at their goals which will result in safety violation. We break this proof into three parts consistent with the three phases shown in Fig. 9(a):

Phase 1 $\to$ Phase 2: Let $t = t_1$ be the time at which phase 1 ends (and phase 2 starts) i.e. when robots fall in deadlock. In theorem 3 we showed that in deadlock $\|\Delta \boldsymbol{p}_{21}\| = D_s$, and in theorem 4 we showed that the positions of robots and their goals are collinear. So at the end of phase 1, $\Delta \boldsymbol{p}_{21}(t_1) = D_s \hat{\boldsymbol{e}}_{\beta+\pi}$. The goal vector $\Delta \boldsymbol{p}_{d_{21}}(t) := \boldsymbol{p}_{d_2} - \boldsymbol{p}_{d_1} = D_G \hat{\boldsymbol{e}}_\beta \; \forall t > 0$.

Phase 2 $\longrightarrow$ Phase 3: The initial condition of phase two is the final condition of phase one i.e. $\Delta \boldsymbol{p}_{21}(t_1) = D_s \hat{\boldsymbol{e}}_{\beta+\pi}$. In phase two, we use feedback linearization to rotate the assembly of robots making sure that the distance between them stays at D_s, until the orientation of the

vector $\Delta \boldsymbol{p}_{21}(t) = D_s \hat{\boldsymbol{e}}_{\theta(t)}$ aligns with $\Delta \boldsymbol{p}_{d_{21}} = D_G \hat{\boldsymbol{e}}_\beta$. (Such a controller is guaranteed to exist (see Grover et al. [2019]). Once done, $\exists$ a time t_2 at which $\theta(t_2) = \beta$. Moreover, at $t = t_2$, the robots are no longer moving and $\Delta \boldsymbol{p}_{21}(t_2) = D_s \hat{\boldsymbol{e}}_\beta$. These states are the final condition for phase 2 and initial condition for phase 3.

Phase 3 $\longrightarrow \infty$: In this phase, the initial condition is $\Delta \boldsymbol{p}_{21}(t_2) = D_s \hat{\boldsymbol{e}}_\beta$. Also, note that the dynamics of phase 3 control are specified by a proportional controller. The dynamics of relative positions and velocities are:

$$\Delta \dot{\boldsymbol{p}}_{21} = -k_p(\Delta \boldsymbol{p}_{21} - \Delta \boldsymbol{p}_{d_{21}}) \tag{90}$$

where $\Delta \boldsymbol{p}_{d_{21}} = D_G \hat{\boldsymbol{e}}_\beta$. Now, we will do a coordinate change as described next. Let $\Delta \tilde{\boldsymbol{p}}_{21} := R_{-\beta} \Delta \boldsymbol{p}_{21}$. The initial conditions in these coordinates are $\Delta \tilde{\boldsymbol{p}}_{21}(t_2) = R_{-\beta} D_s \hat{\boldsymbol{e}}_\beta = (D_s, 0)$ i.e. $\Delta \tilde{p}_{21}^x(t_2) = D_s, \Delta \tilde{p}_{21}^y(t_2) = 0$. The dynamics in new coordinates are:

$$\Delta \dot{\tilde{\boldsymbol{p}}}_{21} = -k_p(\Delta \tilde{\boldsymbol{p}}_{21} - R_{-\beta} \Delta \boldsymbol{p}_{d_{21}}) \tag{91}$$

Using these coordinates, note that $R_{-\beta} \Delta \boldsymbol{p}_{d_{21}} = (D_G, 0)$. Note from the dynamics and the initial conditions for the y components of relative position that the only solution is the zero solution i.e. $\Delta \tilde{p}_{21}^y(t) \equiv 0 \; \forall \; t \geq t_2$. As for the x component, we can compute the solution to be $\Delta \tilde{p}_{21}^x(t) = D_G + (D_s - D_G)e^{-k_p(t-t_2)}$. Finally, note that $\frac{d\|\Delta \boldsymbol{p}_{12}(t)\|}{dt} = -k_p(D_s - D_G)e^{-k_p(t-t_2)} > 0$ (from Assumption 4). Hence, the distance between the robots is non-decreasing *i.e.* the robots never fall in deadlock. Additionally, since the robots use a proportional controller, their positions exponentially stabilize to their goals. $\quad\square$

References

Alonso-Mora, J., DeCastro, J. A., Raman, V., Rus, D. and Kress-Gazit, H. [2018], 'Reactive mission and motion planning with deadlock resolution avoiding dynamic obstacles', *Autonomous Robots* **42**(4), 801–824.

Ames, A. D., Xu, X., Grizzle, J. W. and Tabuada, P. [2017], 'Control barrier function based quadratic programs for safety critical systems', *IEEE Transactions on Automatic Control* **62**(8), 3861–3876.

Borrmann, U., Wang, L., Ames, A. D. and Egerstedt, M. [2015], 'Control barrier certificates for safe swarm behavior', *IFAC-PapersOnLine* **48**(27), 68–73.

Boyd, S. and Vandenberghe, L. [2004], *Convex optimization*, Cambridge university press.

Burgard, W., Moors, M., Stachniss, C. and Schneider, F. E. [2005], 'Coordinated multi-robot exploration', *IEEE Transactions on robotics* **21**(3), 376–386.

Cortes, J., Martinez, S., Karatas, T. and Bullo, F. [2004], 'Coverage control for mobile sensing networks', *IEEE Transactions on robotics and Automation* **20**(2), 243–255.

Douthwaite, J. A. [2019], Multi-agent Collision Avoidance Using Interval Analysis and Symbolic Modelling with its Application to the Novel Polycopter, PhD thesis, University of Sheffield.

Douthwaite, J. A., Zhao, S. and Mihaylova, L. S. [2018], A comparative study of velocity obstacle approaches for multi-agent systems, in '2018 UKACC 12th International Conference on Control (CONTROL)', IEEE, pp. 289–294.

Douthwaite, J. A., Zhao, S. and Mihaylova, L. S. [2019], 'Velocity obstacle approaches for multi-agent collision avoidance', *Unmanned Systems* **7**(01), 55–64.

Duhaut, D., Carrillo, E. and Saint-Aime, S. [2007], Avoiding deadlock in multi-agent systems, *in* '2007 IEEE International Conference on Systems, Man and Cybernetics', pp. 1642–1647.

Feng, S. W. and Yu, J. [2019], 'Optimal perimeter guarding with heterogeneous robot teams: complexity analysis and effective algorithms', *IEEE Robotics and Automation Letters* **5**(2), 430–437.

Fiorini, P. and Shiller, Z. [1998], 'Motion planning in dynamic environments using velocity obstacles', *The International Journal of Robotics Research* **17**(7), 760–772.

Grover, J., Liu, C. and Sycara, K. [2019], 'Deadlock analysis and resolution in multi-robot systems (extended version)'. **URL:** *https://arxiv.org/abs/1911.09146*

Grover, J., Liu, C. and Sycara, K. [2020a], 'Why does symmetry cause deadlocks?', *IFAC-PapersOnLine* **53**(2), 9746–9753.

Grover, J. S., Liu, C. and Sycara, K. [2020b], 'Parameter identification for multirobot systems using optimization based controllers (extended version)', *arXiv preprint arXiv:2009.13817* .

Hokayem, P. F., Stipanović, D. M. and Spong, M. W. [2010], 'Coordination and collision avoidance for lagrangian systems with disturbances', *Applied Mathematics and Computation* **217**(3), 1085–1094.

Jager, M. and Nebel, B. [2001], Decentralized collision avoidance, deadlock detection, and deadlock resolution for multiple mobile robots, *in* 'Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium (Cat. No. 01CH37180)', Vol. 3, IEEE, pp. 1213–1219.

Kantor, G., Singh, S., Peterson, R., Rus, D., Das, A., Kumar, V., Pereira, G. and Spletzer, J. [2003], Distributed search and rescue with robot and sensor teams, *in* 'Field and Service Robotics', Springer, pp. 529–538.

Leitmann, G. and Skowronski, J. [1977], 'Avoidance control', *Journal of optimization theory and applications* **23**(4), 581–591.

Leitmann, G. and Skowronski, J. [1983], 'A note on avoidance control', *Optimal Control Applications and Methods* **4**(4), 335–342.

Li, Y., Gupta, K. and Payandeh, S. [2005], Motion planning of multiple agents in virtual environments using coordination graphs, *in* 'Proceedings of the 2005 IEEE International Conference on Robotics and Automation', IEEE, pp. 378–383.

O'DONNELL, P. [1989], Deadlockfree and collision-free coordination of two robot manipulators., *in* 'Proc. IEEE International Conference on Robotics and Automation', pp. 484–489.

Ogren, P., Egerstedt, M. and Hu, X. [2002], 'A control lyapunov function approach to multiagent coordination', *IEEE Transactions on Robotics and Automation* **18**(5), 847–851.

Olfati-Saber, R., Fax, J. A. and Murray, R. M. [2007], 'Consensus and cooperation in networked multi-agent systems', *Proceedings of the IEEE* **95**(1), 215–233.

Pan, T., Verginis, C. K., Wells, A. M., Kavraki, L. E. and Dimarogonas, D. V. [2020], Augmenting control policies with motion planning for robust and safe multi-robot navigation, *in* '2020 IEEE/RSJ International Conference on Intelligent

Robots and Systems (IROS)', IEEE, pp. 6975–6981.

Petti, S. and Fraichard, T. [2005], Safe motion planning in dynamic environments, *in* '2005 IEEE/RSJ International Conference on Intelligent Robots and Systems', IEEE, pp. 2210–2215.

Rodríguez-Seda, E. J., Stipanović, D. M. and Spong, M. W. [2016], 'Guaranteed collision avoidance for autonomous systems with acceleration constraints and sensing uncertainties', *Journal of Optimization Theory and Applications* **168**(3), 1014–1038.

Stipanović, D. M., Hokayem, P. F., Spong, M. W. and Šiljak, D. D. [2007], 'Cooperative avoidance control for multiagent systems', *Journal of dynamic systems, measurement, and control* **129**(5), 699–707.

Van Den Berg, J., Guy, S. J., Lin, M. and Manocha, D. [2011], Reciprocal n-body collision avoidance, *in* 'Robotics research', Springer, pp. 3–19.

Van den Berg, J., Lin, M. and Manocha, D. [2008], Reciprocal velocity obstacles for real-time multi-agent navigation, *in* '2008 IEEE International Conference on Robotics and Automation', IEEE, pp. 1928–1935.

Wang, L., Ames, A. D. and Egerstedt, M. [2017], 'Safety barrier certificates for collisions-free multirobot systems', *IEEE Transactions on Robotics* **33**(3), 661–674.

Wilkie, D., Van Den Berg, J. and Manocha, D. [2009], Generalized velocity obstacles, *in* '2009 IEEE/RSJ International Conference on Intelligent Robots and Systems', IEEE, pp. 5573–5578.

Yamaguchi, H. [1999], 'A cooperative hunting behavior by mobile-robot troops', *the International Journal of robotics Research* **18**(9), 931–940.

Zhou, Y., Hu, H., Liu, Y. and Ding, Z. [2017a], 'Collision and deadlock avoidance in multirobot systems: A distributed approach', *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **47**(7), 1712–1726.

Zhou, Y., Hu, H., Liu, Y. and Ding, Z. [2017b], 'Collision and deadlock avoidance in multirobot systems: A distributed approach', *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **47**(7), 1712–1726.