Determination of All Stable and Unstable Equilibria for Image-Point-Based Visual Servoing

Alessandro Colotti[1], Jorge García Fontán[2], Alexandre Goldsztejn[1], Sébastien Briot[1], François Chaumette[3], Olivier Kermorgant[1], Mohab Safey El Din[2]

***Abstract*—Local minima are a well-known drawback of image-based visual servoing systems. Up to now, there were no formal guarantees on their number, or even their existence, according to the considered configuration. In this work, a formal approach is presented for the exhaustive computation of all minima and unstable equilibria for a class of six well-known image-based visual servoing controllers. This approach relies on a new polynomial formulation of the equilibrium condition that avoids using the camera pose. By using modern computational algebraic geometry methods and an *ad hoc* symmetry breaking strategy, the formal resolution of this new equilibrium condition is rendered computationally feasible. The proposed methodology is applied to compute the equilibria of several classical visual servoing tasks, with planar and non-planar configurations of four and five points. The effects of local minima and saddle points on the dynamics of the system are finally illustrated through intensive simulation results, as well as the effects of image noise and uncertainties on depths.**

***Index Terms*—Visual Servoing, Formal Methods in Robotics and Automation, Sensor-based Control, Stability Analysis**

I. Introduction

Visual servoing is a mature area whose formalism is well established [1]. Even if it has lead to many successful applications and benefits in practice from a large convergence domain and a large robustness to modeling and calibration errors, its stability analysis is still an open theoretical issue when all the six degrees of freedom of the system have to be controlled.

In this regard, pose-based visual servoing [2], [3] seems to be appealing since using a minimal representation of pose as input of the control scheme allows the system to be globally asymptotically stable (i.e., the system will converge whatever its initial configuration) in case the pose is assumed to be perfectly estimated all along the servo. Without speaking about the visibility constraint that imposes a sufficient number of image measurements are always available to estimate the pose, the assumption of perfect pose estimations is very strong and can be violated, as exhibited in [4] for the case of a simple planar object. This is due to the fact that going from 2D image measurements to pose is an inverse problem whose Jacobian can be ill-conditioned.

The situation is the same for the hybrid and 2 1/2 D approaches [5]–[7] that use once again a minimal number of visual features as inputs of the control scheme, some expressed in the image, others in 3D. In that case, global asymptotic stability can be demonstrated by assuming the 3D data involved are perfectly estimated.

Another approach consists in using as visual features the image coordinates of at least three points and their depth [8], [9], leading to 3D features. It has been demonstrated in [10] that four equilibrium poses exist in that case, one corresponding to the desired pose and the three others corresponding to unstable equilibria, which means that the system will move away from them as soon as noise will be introduced. Very recently, a deeper study of this scheme has been performed in [11], leading to the nice result that the system is almost globally asymptotically stable in perfect conditions, but also that there exists a set of particular poses that will converge towards these unstable equilibria, meaning they correspond to saddle points.

Very few results exist concerning image-based visual servoing. This approach has the strong advantage of being the more robust to measurement noise since the inputs of the closed-loop control scheme only rely on image measurements. Even if depth is involved in the control scheme since it appears in the interaction matrix from which the control scheme is designed, noise on depth will affect the transient phase of the dynamic system (that is, the trajectory to reach the goal will be perturbed by noise or coarse approximation on depth), but it will have no effect on the accuracy reached at the goal (if the system converges to that goal!). However, up-to-now, only local asymptotic stability has been demonstrated for image-based visual servoing [1], which means one is sure that the system converges if the initial configuration is in the neighborhood of the desired one (and determining quantitatively the size of this neighborhood is a clear open issue). This is due to the fact that a redundant number of visual features, typically the coordinates of at least four image points, are used as inputs of the control scheme for avoiding the famous cylinder of singularities when only three image points are considered [12], [13], and also since there usually exist four different poses such that the image of three points is the same [14]. Global asymptotic stability seems to be out of reach since the existence of attractive local minima has been exhibited in [4]. It concerned the case of coplanar points with a desired configuration not parallel to the image plane, following the result given in [15] that pose ambiguity exists in that particular case. More precisely, if a unique pose exists from the perspective projection of four points, two symmetric solutions exist in the non-parallel case with the para-perspective model,

[1] Nantes Université, École Centrale de Nantes, CNRS, LS2N, UMR 6004, Nantes, France
[2] Sorbonne Université, LIP6, Paris, France
[3] Inria, Univ Rennes, CNRS, IRISA, Rennes, France
Corresponding author e-mail: alessandro.colotti@ls2n.fr
This work has been granted by ANR SESAME, No. ANR-18-CE33-0011.

leading to two similar (but not exactly the same) images for the poses corresponding to these two solutions with the perspective model. Designing one solution as the desired pose and starting from a pose near the other solution, it was not surprising that the system converges toward the latter, a local minimum, and not toward the former corresponding to the global minimum and desired pose. Furthermore, in a recent work focused on the investigation of singularities [16], local minima have been incidentally exhibited for the case of four non coplanar points. These minima have been found by chance in simulation while studying the behavior of the system near such singularities.

The main contribution of this paper is a method to determine the full set of equilibrium points of image-based visual servoing using Cartesian coordinates of image points as visual features, that is, the local minima, the local maxima, and the saddle points, in function of the configuration of the object observed. The exhaustive computation of all equilibria, which has been an open issue in the field for more than 25 years, is a first step toward answering fundamental questions like: is there only one local minimum in the case of a planar object non-parallel to the image plane? Are there other equilibrium points than the global minimum in the case of a planar object parallel to the image plane? What is the situation for non planar objects? Furthermore, we believe that the resolution of this problem is also a first step toward a deep stability analysis of these systems. In particular, and as illustrated with typical test cases with four and five points, the computation of saddle points opens the door to determine in the future the convergence domain towards the desired global minimum.

The paper is organized as follows. In Section II, we introduce two classes of image-based visual servoing controllers and we derive polynomial models for their equilibrium conditions. A first polynomial model is naturally expressed in the state-space of the system, but turns out to be intractable. We then propose a second model, expressed in an extended space of visual features, which lead to systems of equations that are medium size and strongly non-linear. In Section III, we show that computational algebraic geometry methods are suitable to solve this second model. Our solver `msolve` [17] is a recent, highly optimized implementation based on Gröbner bases computations, which is presented in Section III-A. A modification of the polynomial model dedicated to an efficient resolution is presented in Section III-B. This allows us to compute all fixed points for a benchmark of typical configurations of image-based visual servoing systems, presented in Section IV. The application of our strategy uncovered a great diversity in the zoology of equilibrium points, a diversity that is further highlighted through extensive simulations for each configuration. Resolution timings ranging from fractions of a second for simple configurations and up to 35 hours for general ones on our 12 core server demonstrate the difficulty of solving these problems. The validity of our results in the presence of realistic perturbations is then examined in Section V, where we present a series of simulations to illustrate that the qualitative behavior of the system remains basically unaffected by the presence of noise.

II. Model derivation for the equilibria's computation

A. IBVS Model & Control

We consider a traditional IBVS task [1], where we take the image-plane coordinates of $N \geq 4$ points (not necessarily coplanar) as visual features. We assume that we have an ideal *pinhole camera*, that the points are visible from any direction and that the camera is controllable in velocity. The points' Cartesian coordinates ${}^o\mathbf{a}_i$, expressed in a given world frame, are constant and known *a priori* for the analysis. The visual features x_i, y_i of the i-th point ${}^c\mathbf{a}_i = (X_i, Y_i, Z_i)$, expressed in the camera frame, are:

$$x_i = \frac{X_i}{Z_i}\ ,\ y_i = \frac{Y_i}{Z_i} \tag{1}$$

With this choice of features, the interaction matrix $\mathbf{L}_i$ related to $\mathbf{s}_i = (x_i, y_i)$ is given by [1]:

$$\mathbf{L}_i = \begin{bmatrix} -\frac{1}{Z_i} & 0 & \frac{x_i}{Z_i} & x_i y_i & -(1+x_i^2) & y_i \\ 0 & -\frac{1}{Z_i} & \frac{y_i}{Z_i} & 1+y_i^2 & -x_i y_i & -x_i \end{bmatrix} \tag{2}$$

Denoting with $\mathbf{s} = (\mathbf{s}_1, \ldots, \mathbf{s}_N)$ the (stacked) image features and with $\mathbf{s}^* = (\mathbf{s}_1^*, \ldots, \mathbf{s}_N^*)$ the desired ones, the dynamics of the features error $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ are:

$$\dot{\mathbf{e}} = \mathbf{L}_\mathbf{e}\boldsymbol{\tau}_c, \qquad \mathbf{L}_\mathbf{e} = [\mathbf{L}_1^T, \ldots, \mathbf{L}_N^T]^T, \tag{3}$$

where $\boldsymbol{\tau}_c = (\mathbf{v}_c, \boldsymbol{\omega}_c)$ is the camera's spatial velocity that represents the control inputs, with $\mathbf{v}_c$ being the linear velocity and $\boldsymbol{\omega}_c$ being the angular one.

The general goal of an IBVS task is to make the image features $\mathbf{s}$ converge to the desired features $\mathbf{s}^*$. Several strategies to design appropriate controllers for these tasks exist in the literature. We focus in this work on three such strategies [18], [19], that are all based on the general form:

$$\boldsymbol{\tau}_c = -\lambda \mathbf{C}_\mathbf{e}\mathbf{e} \text{ with } \lambda \in \mathbb{R}_+. \tag{4}$$

More precisely, we consider:

- *Transpose* controller: $\mathbf{C}_\mathbf{e} = \widehat{\mathbf{L}}_\mathbf{e}^T$,
- *Pseudo-inverse* controller: $\mathbf{C}_\mathbf{e} = \widehat{\mathbf{L}}_\mathbf{e}^+$,
- *Levenberg-Marquardt* controller: $\mathbf{C}_\mathbf{e} = (\widehat{\mathbf{L}}_\mathbf{e}^T\widehat{\mathbf{L}}_\mathbf{e} + \mu\mathbf{I})^{-1}\widehat{\mathbf{L}}_\mathbf{e}^T$, with $\mu \in \mathbb{R}_+$.

We denote with $\widehat{\mathbf{L}}_\mathbf{e}$ the interaction matrix used in the control feedback, which might be different from the true interaction matrix $\mathbf{L}_\mathbf{e}$, since the latter depends on the features depths that may not be available in practice. In particular, in this work we consider two classical choices for $\widehat{\mathbf{L}}_\mathbf{e}$, which give rise to the two following classes of control schemes:

- *Perfect approximation*: we assume that the image features $\mathbf{s}$ and the points' depths Z_i are perfectly known at each moment, with no approximation nor noise. The true interaction matrix $\mathbf{L}_\mathbf{e}$ is then used in the control input, i.e. $\widehat{\mathbf{L}}_\mathbf{e} = \mathbf{L}_\mathbf{e}$.
- *Desired pose approximation*: we assume that the depths Z_i are unknown. The true interaction matrix is then approximated as $\widehat{\mathbf{L}}_\mathbf{e} = \mathbf{L}_{\mathbf{e}^*}$, where $\mathbf{L}_{\mathbf{e}^*}$ is the value of $\mathbf{L}_\mathbf{e}$ with desired features $\mathbf{s} = \mathbf{s}^*$ and desired depths Z_i^*.

The main objective of this paper is the computation of all equilibria for these six controllers. Before proceeding, the following lemma shows how the equilibria characterization for all strategies can be actually reduced to the case of the transpose controller.

Lemma 1. *Given the interaction matrix $\widehat{\mathbf{L}}_\mathbf{e}$ to be used in the control loop, the system's equilibrium configurations, characterized by:*

$$\widehat{\mathbf{L}}_\mathbf{e}^T\mathbf{e} = 0, \tag{5}$$

are the same for the pseudo-inverse, transpose and Levenberg-Marquardt controllers.

Proof. We argue that $\boldsymbol{\tau}_c = 0$ if and only if $\widehat{\mathbf{L}}_\mathbf{e}^T\mathbf{e} = 0$ for all three controllers. This is obvious for the transpose controller. It is well known [20] that $\ker \widehat{\mathbf{L}}_\mathbf{e}^+ = \ker \widehat{\mathbf{L}}_\mathbf{e}^T$, proving the equivalence for the pseudo-inverse controller. Finally, by the positive semi-definiteness of $\mathbf{L}_\mathbf{e}^T\mathbf{L}_\mathbf{e}$, $(\mathbf{L}_\mathbf{e}^T\mathbf{L}_\mathbf{e} + \mu\mathbf{I})$ is invertible for all $\mu > 0$, meaning that $(\mathbf{L}_\mathbf{e}^T\mathbf{L}_\mathbf{e} + \mu\mathbf{I})^{-1}\mathbf{L}_\mathbf{e}^T\mathbf{e} = 0$ if and only if $\widehat{\mathbf{L}}_\mathbf{e}^T\mathbf{e} = 0$, concluding the proof. □

Furthermore, as it is classically done, we assess the stability properties of a given equilibrium by computing the system's linearization at it and verifying the signs of the linearized system's eigenvalues (see, e.g., [21]). It is well known that, if there are no eigenvalues with null real part, then the linearized and non-linear systems are equivalent in a neighborhood of the equilibrium, and the number of positive eigenvalues (called the equilibrium's *index*) identifies the *type* of equilibrium. More precisely, if the index is null, the equilibrium is stable and corresponds to a minimum while it is an unstable maximum if the index is 6 and a saddle otherwise. In practice, this analysis boils down to computing the vector field's Jacobian in state-space at the equilibrium points. This Jacobian is defined as the derivative of the control input $\boldsymbol{\tau}_c$ with respect to the camera pose. From (4), this Jacobian is given by $-\lambda\mathbf{C}_\mathbf{e}\mathbf{L}_\mathbf{e}$ in the desired pose approximation case since $\mathbf{C}_\mathbf{e}$ is constant in that case and we have $\dot{\boldsymbol{\tau}}_c = -\lambda\mathbf{C}_\mathbf{e}\dot{\mathbf{e}} = -\lambda\mathbf{C}_\mathbf{e}\mathbf{L}_\mathbf{e}\boldsymbol{\tau}_c$. It has the same form in the perfect approximation case for the global minimum where $\mathbf{s} = \mathbf{s}^*$ since $\dot{\mathbf{C}}_\mathbf{e}\mathbf{e} = 0$, but has a more complex form, not given here, for the other equilibria.

The explicit, polynomized expression of (5) is given in (6) (Table I), where $\hat{x}_i$, $\hat{y}_i$ and $\hat{Z}_i$ denote the features used in the evaluation of $\widehat{\mathbf{L}}_\mathbf{e}$. In (6), we multiplied the first three equations by the product of all Z_i's to make the expression polynomial, which is a necessary step to be able to solve the system using computational algebra methods. However, (6) is expressed in what is called the *extended features* $\boldsymbol{\xi} = (\mathbf{s}, \mathbf{Z}) \in \mathbb{R}^{3N}$ [10], meaning that it only provides a necessary condition for the equilibrium: it is a system of 6 equations in $3N > 6$ unknowns, where *infeasible extended features* (i.e., extended features not corresponding to any camera pose) are equally considered together with *feasible extended features*. Thus, it is imperative to find a strategy that allows us to impose that the solutions of (6) are within the set of feasible extended features.

For the sake of clarity, we focus in the next two subsections on the perfect approximation case. The desired pose controllers can then be seen as a (significantly simpler) particular case, which is briefly discussed in Section II-D.

TABLE I
EQUILIBRIUM CONDITION'S EXPLICIT EXPRESSION.

$$\sum_{i=1}^{N}\left[\left(\prod_{j\neq i}^{N}\hat{Z}_j\right)(x_i^* - x_i)\right] = 0 \tag{6a}$$

$$\sum_{i=1}^{N}\left[\left(\prod_{j\neq i}^{N}\hat{Z}_j\right)(y_i^* - y_i)\right] = 0 \tag{6b}$$

$$\sum_{i=1}^{N}\left[\left(\prod_{j\neq i}^{N}\hat{Z}_j\right)\left(\hat{x}_i(x_i - x_i^*) + \hat{y}_i(y_i - y_i^*)\right)\right] = 0 \tag{6c}$$

$$\sum_{i=1}^{N}\hat{x}_i\hat{y}_i(x_i - x_i^*) + (1 + \hat{y}_i^2)(y_i - y_i^*) = 0 \tag{6d}$$

$$\sum_{i=1}^{N}(1 + \hat{x}_i^2)(x_i^* - x_i) + \hat{x}_i\hat{y}_i(y_i^* - y_i) = 0 \tag{6e}$$

$$\sum_{i=1}^{N}(x_i y_i^* - y_i x_i^*) = 0. \tag{6f}$$

Perfect approximation: $\hat{x}_i,\, \hat{y}_i,\, \hat{Z}_i = x_i,\, y_i,\, Z_i$ (7a)

Desired approximation: $\hat{x}_i,\, \hat{y}_i,\, \hat{Z}_i = x_i^*,\, y_i^*,\, Z_i^*$ (7b)

B. Camera state-space model

A first attempt to solve the equilibrium condition is to consider the camera pose ${}^c\mathbf{T}_o \in \mathrm{SE}(3)$ as the system's variables, from which features $\mathbf{s}({}^c\mathbf{T}_o)$ and their corresponding depths $\mathbf{Z}({}^c\mathbf{T}_o)$ are easily computed, leading to intrinsically feasible extended features that can be substituted in (6). However, $\mathrm{SE}(3)$ does not admit a unique representation. We focus here on two of these representations, namely, rotation matrices and quaternions. We then rewrite (6), which is naturally expressed as a function of $\boldsymbol{\xi}({}^c\mathbf{T}_o)$, for the two parameterizations. First, considering the parameterization ${}^c\mathbf{T}_o = ({}^c\mathbf{R}_o, {}^c\mathbf{t}_o)$, we obtain the points coordinates in camera frame by

$${}^c\mathbf{a}_i = {}^c\mathbf{R}_o\, {}^o\mathbf{a}_i + {}^c\mathbf{t}_o, \tag{8}$$

from which the image features can be retrieved using (1) and their depth from the third component of ${}^c\mathbf{a}_i$. Second, this relation also allows us to rewrite (6) as a function of the quaternion parameterization ${}^c\mathbf{T}_o = ({}^c\mathbf{q}_o, {}^c\mathbf{t}_o)$, by expressing ${}^c\mathbf{R}_o$ as a function of ${}^c\mathbf{q}_o$ (see, e.g., [22]).

In order to actually solve the system on $\mathrm{SE}(3)$, it is additionally necessary to impose the group constraints ${}^c\mathbf{R}_o \in \mathrm{SO}(3)$ or ${}^c\mathbf{q}_o \in \mathbb{H}$ for the two representations. Putting all together, the explicit expression of (6) as a function of the camera state becomes a huge degree $3N$ polynomial system of 12 equations in 12 variables with the $({}^c\mathbf{R}_o, {}^c\mathbf{t}_o)$-representation, while we have a degree $6N$ polynomial system of 7 equations in 7 variables with the $({}^c\mathbf{q}_o, {}^c\mathbf{t}_o)$ representation. A comparison between the degrees of these systems is detailed in Table II. Because of their substantial density and high degrees, the latter

TABLE II
COMPLEXITY COMPARISON BETWEEN DIFFERENT CHOICES OF PARAMETRIZATION FOR THE PERFECT APPROXIMATION CASE.

N	Parametrization	# of unknowns	# of equations	Polynomial degrees of the equations
N	$({}^c\mathbf{R}_o, {}^c\mathbf{t}_o)$	12	12	$[2N-1, 2N-1, 3N-1, 3N, 3N, 2N-1, 2, \overset{(6)}{\ldots}, 2]$
	$({}^c\mathbf{q}_o, {}^c\mathbf{t}_o)$	7	7	$[2(2N-1), 2(2N-1), 2(3N-1), 6N, 6N, 2(2N-1), 2]$
	$\boldsymbol{\xi}$	$3N$	$6 + \frac{N(N-1)}{2} + 1$	$[N, N, N+1, 3, 3, 1, 4, \overset{N(N-1)/2}{\ldots}, 4, 5]$
$N=4$	$({}^c\mathbf{R}_o, {}^c\mathbf{t}_o)$	12	12	$[7, 7, 11, 12, 12, 7, 2, \overset{(6)}{\ldots}, 2]$
	$({}^c\mathbf{q}_o, {}^c\mathbf{t}_o)$	7	7	$[14, 14, 22, 24, 24, 14, 2]$
	$\boldsymbol{\xi}$	12	13	$[4, 4, 5, 3, 3, 1, 4, \overset{(6)}{\ldots}, 4, 5]$
$N=5$	$({}^c\mathbf{R}_o, {}^c\mathbf{t}_o)$	12	12	$[9, 9, 14, 15, 15, 9, 2, \overset{(6)}{\ldots}, 2]$
	$({}^c\mathbf{q}_o, {}^c\mathbf{t}_o)$	7	7	$[18, 18, 28, 30, 30, 18, 2]$
	$\boldsymbol{\xi}$	15	17	$[5, 5, 6, 3, 3, 1, 4, \overset{(10)}{\ldots}, 4, 5]$
$N=6$	$({}^c\mathbf{R}_o, {}^c\mathbf{t}_o)$	12	12	$[11, 11, 17, 18, 18, 11, 2, \overset{(6)}{\ldots}, 2]$
	$({}^c\mathbf{q}_o, {}^c\mathbf{t}_o)$	7	7	$[22, 22, 34, 36, 36, 22, 2]$
	$\boldsymbol{\xi}$	18	22	$[6, 6, 7, 3, 3, 1, 4, \overset{(15)}{\ldots}, 4, 5]$

of which form the basis for determining the maximum number of complex solutions (the so-called Bézout bound), these systems are out of reach for current state-of-the-art complete solvers for polynomial systems, as these solvers' performances are sensitive to these factors. Note that similar conclusions are obtained using minimal representations of rotations, such as the angle multiplied by the unitary rotation axis.

C. Extended features space model

1) Derivation of the model: As explained just above, expressing the system directly in $\mathrm{SE}(3)$ leads to a highly complicated system, which cannot be solved as-is. The equilibria condition (6), considered directly in the extended features space, is seemingly far simpler than the system's expression in camera state-space. We aim here to replace the extended feature feasibility, expressed in the previous system using the camera pose, by simpler constraints acting only on the extended features.

To this end, we propose to check the feasibility of the extended features $\boldsymbol{\xi}$ by first reconstructing the corresponding points in the camera frame by using the one-to-one correspondence

$$Z_i = Z_i \ , \ X_i = x_i Z_i \ , \ Y_i = y_i Z_i, \tag{9}$$

and second imposing that point-to-point inter-distances in the camera frame to be the same as in the world frame. This leads to the following $C_2^N = \frac{1}{2}N(N-1)$ point-to-point inter-distance constraints:

$$(x_i Z_i - x_j Z_j)^2 + (y_i Z_i - y_j Z_j)^2 + (Z_i - Z_j)^2 = d_{ij}^2, \tag{10a}$$

for all $i < j = 1, \ldots, N$, where the distances $d_{ij} := \|{}^o\mathbf{a}_i - {}^o\mathbf{a}_j\|$ are known *a priori*. The new system (6) and (10a), which allows us to compute the system's equilibria directly in extended features space, has more equations but with much lower degrees with respect to the camera state-space representation (see Table II) and it is in the scope of current complete solvers for polynomial systems, as shown in Section III.

This new expression offers a great simplification, but also comes with a drawback: the point-to-point inter-distance constraints are only a necessary condition that is not sufficient for an extended feature to be feasible. Informally, points in the camera frame that have the same point-to-point inter-distance as in the world frame may come from an improper rigid transformation belonging to the group of isometries $\mathrm{E}(3)$ and not to $\mathrm{SE}(3)$.

It is possible to overcome this issue by noting that the handedness of a tetrahedron is only preserved under *proper* isometries. Therefore, the following constraint is satisfied only by *feasible* extended features:

$$\det \begin{bmatrix} x_1 Z_1 & x_2 Z_2 & x_3 Z_3 & x_4 Z_4 \\ y_1 Z_1 & y_2 Z_2 & y_3 Z_3 & y_4 Z_4 \\ Z_1 & Z_2 & Z_3 & Z_4 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \det \begin{bmatrix} {}^o\mathbf{a}_1 & \cdots & {}^o\mathbf{a}_4 \\ 1 & \cdots & 1 \end{bmatrix}, \tag{10b}$$

where $\{{}^o\mathbf{a}_i\}_{i=1}^4$ are 4 points in the configuration and where the two determinants are proportional to the *signed volume* of the tetrahedron (see, e.g., [23]) represented in the camera-frame and world-frame, respectively.

In the non-coplanar case, by taking 4 non-coplanar points in (10b), the *feasibility constraints* (10) select only the feasible extended features. In the coplanar case, handedness

loses meaning due to the tetrahedron's degeneracy and the distance constraints (10a) become sufficient for feasibility. Nevertheless, the addition of (10b) still has the benefit of speeding up the computations: informally, extended features corresponding to proper and improper isometries are merged in the coplanar case, leading to solutions of multiplicity two, and the determinant constraint removes this multiplicity.

2) Pose reconstruction: We discuss here a classical pose reconstruction strategy to recover the camera pose given the extended features $\boldsymbol{\xi}$ and the points' coordinates in world frame $\{{}^o\mathbf{a}_i\}_{i=1}^N$. The proposed reconstruction algorithm is based on the solution of the *orthogonal Procrustes' problem* [24], whose aim, given two sets of points $\{\mathbf{P}_i\}_{i=1}^N$ and $\{\mathbf{Q}\}_{i=1}^N$, is to find an orthogonal matrix $\boldsymbol{\Omega}$ that maps $\mathbf{P}_i$ to $\mathbf{Q}_i$ for all i in the best possible way. The problem allows for the points to undergo a translation as well; in this case, the translation is first removed from the problem by shifting the points' centroid to the origin, and then reconstructed using $\boldsymbol{\Omega}$ and the centroids' information.

The reconstruction algorithm for a given solution $\boldsymbol{\xi}$ of (6) and (10) is detailed in Table III. The feasibility of $\boldsymbol{\xi}$ implies that there always exists at least one proper camera pose that corresponds to it, or, equivalently, that there exists at least one proper isometry between the two sets of points $\mathbf{P}$ and $\mathbf{Q}$. Such isometry is a global minimizer of the residual $\|\boldsymbol{\Omega}\mathbf{P}-\mathbf{Q}\|_F$, meaning that it is retrieved by the reconstruction algorithm, being the reconstructed orientation guaranteed to minimize the residual (see [24]–[26]). However, since the algorithm computes only one camera pose, the uniqueness of such solution is critical to ensure that no equilibria are lost due to the reconstruction process. Said differently, we want to prove that there exists a one-to-one correspondence between feasible extended features and proper camera poses.

In the non-coplanar case, it is possible to show[1] that the (proper) solution $\boldsymbol{\Omega}$ computed in Step 3 is unique. Similarly, when the points are coplanar (but non-collinear), one can show[2] that there are exactly one proper and one improper solution. We can then select the unique proper rotation by forcing its determinant to be positive (see, e.g., [26]). Thus, in both the non-coplanar and the coplanar cases, the proposed reconstruction algorithm always finds the unique corresponding camera pose for all solutions of (6) and (10).

D. Differences between perfect and desired pose approximation controllers

In the desired pose controller case, the extended features space's model (6) and (10) still represents the easiest way to compute all system's equilibria. By substituting $\hat{x}_i$, $\hat{y}_i$ and $\hat{Z}_i$ with the desired x_i^*, y_i^* and Z_i^*, the equilibrium condition (6) becomes a linear system of equations with respect to the extended features $\boldsymbol{\xi}$, which significantly reduces the complexity of its resolution, while the feasibility constraints (10) are not affected. Once all solutions are computed, it is then possible to retrieve the camera pose using the reconstruction strategy presented in Section II-C2.

TABLE III
POSE RECONSTRUCTION ALGORITHM.

Input: $\{{}^o\mathbf{a}_i\}_{i=1}^N$ and $\boldsymbol{\xi}$ which is a solution of (6) and (10)
Output: $({}^c\mathbf{R}_o, {}^c\mathbf{t}_o)$

1. Determine, through (9), $\{{}^c\mathbf{a}_i\}_{i=1}^N$ and define $\mathbf{Q}_i \coloneqq {}^c\mathbf{a}_i - {}^c\mathbf{k}$ and $\mathbf{P}_i \coloneqq {}^o\mathbf{a}_i - {}^o\mathbf{k}$, where ${}^c\mathbf{k}$ and ${}^o\mathbf{k}$ represent the points' centroid in camera- and world-frame, respectively.
2. Let $\mathbf{Q}$ and $\mathbf{P}$ be the $3 \times N$ matrices that have the corresponding points as columns, define the *covariance matrix* $\mathbf{M} = \mathbf{Q}\mathbf{P}^T$ and let $\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ be its SVD decomposition.
3. The orientation-preserving rotation matrix is $\boldsymbol{\Omega} = \mathbf{U}\,\mathrm{diag}(1,1,d)\,\mathbf{V}^T$, where d is the sign of the determinant of $\mathbf{U}\mathbf{V}^T$.
4. Return the reconstructed camera pose as $({}^c\mathbf{R}_o,\, {}^c\mathbf{t}_o) = (\boldsymbol{\Omega},\, {}^c\mathbf{k} - \boldsymbol{\Omega}\,{}^o\mathbf{k})$.

III. EQUILIBRIA'S COMPUTATION

This section deals with the resolution of the polynomial systems derived in the previous sections by means of computational algebraic geometry methods and, in particular, Gröbner bases computations.

A. Generalities on computational geometry

The tools used to solve polynomial systems fall into two categories: numerical (e.g. Newton's method, numerical homotopy continuation) [28]–[30], and symbolic (e.g. multivariate resultant, Gröbner bases) [31], [32]. Numerical methods are generally not successful in finding all solutions in an unbounded domain with certification, see, e.g., a comparison to homotopy methods in [33]. Symbolic methods are thus privileged for applications where an exhaustive search of all solutions is required.

Given a set of polynomials $f_1, \dots, f_s$ in $\mathbb{Q}[\mathbf{x}]$, the ring of polynomials in variables $\mathbf{x} = (x_1, \dots, x_n)$ with rational coefficients, we say that they generate a *polynomial ideal*, i.e. the set defined by all the polynomials that are algebraic combinations of $f_1, \dots, f_s$:

$$I = \langle f_1, \dots, f_s \rangle = \left\{ \sum_{i=1}^{s} h_i f_i \mid h_i \in \mathbb{Q}[x_1, \dots, x_n] \right\}. \quad (11)$$

For any other polynomial g in the ideal I, if a point $\mathbf{c} \in \mathbb{C}^n$ satisfies $f_1(\mathbf{c}) = \cdots = f_s(\mathbf{c}) = 0$, then $g(\mathbf{c}) = 0$ too. This shows that the ideal I is an object to study in order to obtain information about the solutions of the system $f_1(\mathbf{x}) = \cdots = f_s(\mathbf{x}) = 0$.

The locus of all the *complex* solutions of this system is called the *algebraic variety* defined by the ideal I, denoted

$$\mathbf{V}(I) = \{\mathbf{c} \in \mathbb{C}^n \mid f(\mathbf{c}) = 0,\ \forall f \in I\}. \quad (12)$$

[1] A global minimizer $\boldsymbol{\Omega}^*$ of the residual satisfies $\boldsymbol{\Omega}^* = \mathbf{U}\mathbf{S}\mathbf{V}^T$, where $\mathbf{S} \in \mathrm{O}(3)$ is a maximizer of $\langle \mathbf{S}, \boldsymbol{\Sigma} \rangle_F$ and $\mathbf{Q}\mathbf{P}^T = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ (see [27, p. 327]). When the points are non-coplanar, the singular values of Σ are strictly positive, which implies that $\mathbf{S} = \mathbf{I}$ is the unique maximizer of $\langle \mathbf{S}, \boldsymbol{\Sigma} \rangle_F$. Since $\boldsymbol{\xi}$ satisfies (10b), then $\mathrm{sign}\left[\det \mathbf{U}\mathbf{V}^T\right] = 1$, meaning that Step 3 always selects $\mathbf{S} = \mathbf{I}$.

[2] When the points are coplanar but non-collinear, by standard linear algebra we have that $\boldsymbol{\Sigma}$ has exactly one null singular value, from which it directly follows that $\langle \mathbf{S}, \boldsymbol{\Sigma} \rangle_F$ has exactly two maximizers, which are $\mathbf{S}_1 = \mathbf{I}_3$ and $\mathbf{S}_2 = \mathrm{diag}\{1, 1, -1\}$. We then have $\det \mathbf{U}\mathbf{S}_1\mathbf{V}^T = -\det \mathbf{U}\mathbf{S}_2\mathbf{V}^T$, showing that one solution is proper while the other is improper.

When the algebraic variety of an ideal consists of finitely many (complex) roots, we say that the ideal is *zero-dimensional* (because its variety is a set of dimension zero). In this case, we define the *degree* of the ideal as the number of complex roots of the system, counted with multiplicity. We refer to [34] for more information about polynomial ideals and varieties.

Gröbner bases are an essential tool in computer algebra for solving problems with polynomial ideals. They can be used, among other things, whether a polynomial is contained in a given ideal, to eliminate a subset of variables from a system of equations, or to obtain a rational parametrization of the solutions. Given a set of polynomials defining an ideal and given a monomial ordering (i.e. a hierarchy of the monomials in the polynomial ring), which must be specified *a priori*, a Gröbner basis is a set of generators of that ideal that have particularly useful computational properties. Different monomial orderings have different effects on the structure of the final Gröbner bases and on the complexity of their computation. For instance, the *lexicographical* ordering provides a triangular description of the ideal, but it is costly to compute, while the *degree reverse lexicographical* ordering results in general in the fastest computations and in polynomials of lower degree. We refer to [32], [34] for a more complete introduction to Gröbner bases and their applications.

The current state-of-the-art methods for computing Gröbner bases are based on the F4 [35] and F5 [36] algorithms. The FGLM algorithm [37] for change of ordering allows to compute a lexicographical ordering from a given Gröbner basis with respect to an *easier* monomial ordering. Efficient implementations of these algorithms for solving systems of polynomials exist in computer algebra systems like Magma or Maple, or in libraries such as FGb [38] or `msolve` [17]. In this paper we relied on `msolve`, an open-source, high-performance library for computing the real roots of zero-dimensional polynomial systems, which is based on efficient implementations of F4 and FGLM and on a dedicated univariate real root isolation algorithm.

Using `msolve`, all complex solutions of the system (6) and (10) are computed together with their multiplicity, real solutions being then easily selected. Note that this system is overconstrained; while this may be a problem for many numerical polynomial solving methods, Gröbner bases computations can deal with it nicely, since the extra relations belonging to the same ideal provide more information about its algebraic structure. Nevertheless, the system (6) and (10) is still difficult to solve for `msolve`, and we were able to improve the resolution process by breaking a symmetry present in the coplanar case through a change of coordinates invariant to this symmetry.

B. Symmetry breaking in the coplanar case

In the coplanar case, the solutions of equations (6) and (10) present a symmetry with respect to a reflection of the scene through the optical center of the camera[3]. In this section we explain how to exploit this symmetry to our advantage by applying a change of coordinates invariant to this reflection. By means of Gröbner bases and algebraic elimination, we derive a new system of equations in the new variables that is computationally easier to solve.

In particular, consider the transformation that maps the observed points coordinates to their reflection through the camera center by the mapping ${}^c\mathbf{a}_i \mapsto -{}^c\mathbf{a}_i$. Each point with coordinates $(\mathbf{s}, \mathbf{Z})$ in the space of the extended features is mapped to $(\mathbf{s}, -\mathbf{Z})$ by this reflection. Applying this transformation to the system (6) and (10) results in a sign change for the polynomials (6a-c), while the remaining equations are left unchanged. The solution set is therefore invariant, i.e., if a point $\boldsymbol{\xi} = (\mathbf{s}, \mathbf{Z})$ is a solution, then so is $\boldsymbol{\xi}' = (\mathbf{s}, -\mathbf{Z})$.

Consider now the following change of variables from the Z_i coordinates to a new set of variables θ_{ij}, which is invariant to the reflection described above

$$\mathbf{Z} = (Z_1, \dots, Z_N) \mapsto \boldsymbol{\theta} = \left(\theta_{12}, \theta_{13}, \dots, \theta_{(N-1)N}\right), \quad (13)$$

with:

$$\theta_{ij} = Z_i Z_j. \quad (14)$$

By applying this coordinate transformation to (6) and (10), we expect to remove the spurious symmetric solutions, and derive a system with half the total number of solutions. Applying the change of variables by hand is extremely hard, but it is possible using Gröbner bases and algebraic elimination [34].

From the definition of θ_{ij} we trivially obtain relations between the old and new variables of the form $\theta_{ij} - Z_i Z_j = 0$. For every polynomial f_i in the original system (6) and (10), we consider the ideal formed by appending these relations, along with a constraint $1 - \ell\, Z_1 \cdots Z_N = 0$, with ℓ an auxiliary variable, to enforce $Z_i \neq 0$ for all i. This ideal lies in the ring $\mathbb{Q}[\ell, \mathbf{s}, \mathbf{Z}, \boldsymbol{\theta}]$:

$$I_i = \langle f_i,\ 1 - \ell\ Z_1 \dots Z_N,\ \theta_{12} - Z_1 Z_2, \theta_{13} - Z_1 Z_3, \dots \rangle. \quad (15)$$

The ideal I_i is then projected onto the space of variables $(\mathbf{s}, \boldsymbol{\theta})$ by computing a Gröbner basis with respect to an elimination ordering, that is, a monomial ordering which eliminates the variables $\mathbf{Z}$ and ℓ:

$$G_i \ \leftarrow\ I_i \cap \mathbb{Q}[\mathbf{s}, \boldsymbol{\theta}]. \quad (16)$$

The basis G_i spans all the algebraic combinations of the polynomials in (15) that only involve $\mathbf{s}$ and $\boldsymbol{\theta}$. It is equivalent to replacing the variables $\mathbf{Z}$ by $\boldsymbol{\theta}$ in the polynomial f_i. However, the resulting G_i consists not of a single polynomial, but rather of a set of polynomials.

Let us illustrate the algebraic elimination step with an example. Consider the case $N = 4$ and let f_1 be the first polynomial (6a). Then, by computing a Gröbner basis for the ideal (15) with respect to the elimination ordering we obtain the following six elements:

$$\begin{aligned} G_1 = [&\theta_{12}\,\theta_{13}\,(x_4 - x_4^\star) + \theta_{12}\,\theta_{14}\,(x_3 - x_3^\star) + \dots, \\ &\theta_{12}\,\theta_{23}\,(x_4 - x_4^\star) + \theta_{12}\,\theta_{24}\,(x_3 - x_3^\star) + \dots, \\ &\theta_{13}\,\theta_{23}\,(x_4 - x_4^\star) + \theta_{12}\,\theta_{34}\,(x_3 - x_3^\star) + \dots, \\ &\theta_{12}\,\theta_{34}\,(x_4 - x_4^\star) + \theta_{14}\,\theta_{24}\,(x_3 - x_3^\star) + \dots, \\ &-\theta_{12}\,\theta_{34} + \theta_{14}\,\theta_{23}, \\ &-\theta_{12}\,\theta_{34} + \theta_{13}\,\theta_{24}], \end{aligned} \quad (17)$$

[3]Note that this is not the case for non-coplanar configurations since this reflection changes the *handedness* of the object, which is preserved by (10b).

where the first four ones are respectively obtained by multiplying f_1 with Z_i, $i = 1, \ldots, 4$, while the last two equations preserve consistency with the change of variables itself, as we can clearly see by substituting (14) in the formulas.

The set union of all the Gröbner bases G_i computed in this form for all the polynomials f_i in (6) and (10) defines the new system of equations in variables $(\mathbf{s}, \boldsymbol{\theta})$. Once again, a constraint of the form $1 - \ell\, \theta_{12} \ldots \theta_{1N} = 0$ must be added to remove the solutions with coordinates $\theta_{ij} = 0$ for some i and j, which correspond to solutions with some coordinate $Z_i = 0$. The new system is then

$$\mathbf{G}(\mathbf{s}, \boldsymbol{\theta}) = (G_1, G_2, \ldots, G_s, 1 - \ell\, \theta_{12} \ldots \theta_{1N}). \tag{18}$$

The system contains more equations and variables than the original one, but the number of complex solutions (i.e. the degree of the ideal) is cut by half, and we observe an improvement in the computation time of up to an order of magnitude (see Section IV). For instance, for $N = 4$, the original system contains 13 equations and 12 variables, while the system (18) consists of 59 equations in 14 variables.

Every solution $(\mathbf{s}, \boldsymbol{\theta})$ for (18) maps to two points $(\mathbf{s}, \mathbf{Z})$ and $(\mathbf{s}, -\mathbf{Z})$, but we are only interested in those for which all the 3D points are in the semi-space in front of the camera (i.e., $Z_i > 0$ for all i). One can easily verify that a solution with positive Z_i exists if and only if all θ_{ij} are positive. Using (14), the depths are then recovered by

$$Z_i = \sqrt{\frac{\theta_{ij}\theta_{ik}}{\theta_{jk}}}, \tag{19}$$

where j and k are chosen arbitrarily such that $i \neq j \neq k$. The pose reconstruction algorithm in Table III is then used to determine the proper camera poses corresponding to these solutions.

Overall, the strategy detailed above allow halving the degree of the polynomial ideal by projecting it onto the space of new variables that are symmetry invariant. This leads to a significant reduction in the computation times, resulting in a more effective formulation for computing the critical points of image point-based visual servoing systems.

IV. Representative test cases

In this section, we apply the equilibria computation strategy developed throughout the paper (of which it is possible to find a summary in Fig. 1) to a selection of case studies. For each of the examples, we compute all the equilibria[4] and we analyze their local stability by looking at the eigenvalues of the vector field's Jacobian, as described in Section II-A. Additionally, we perform extensive simulations for each case by focusing on the pseudo-inverse controller since it is well known to provide a better time response. For a deeper comparison with the transpose controller's performances, and their meaning with respect to the Levenberg-Marquardt one, the interested reader can refer to [39].

[4]The equations given to `msolve` for each case study as well as the list of all computed equilibria are available at https://github.com/acolotti/equilibria-computation-in-ibvs.

The computations of the equilibria are performed on a computing server equipped with an Intel(R) Xeon(R) Gold 6246R CPU running at 3.40GHz, and with 1.5 TB of total memory. We consider examples of *perfect approximation* and *desired pose approximation* control schemes separately in the next two subsections. For the perfect approximation case, we only consider cases having $N = 4$ points, since cases with $N > 4$ are too complex to be tackled by the current version of our solver. For the desired pose approximation, on the other hand, we also consider one case having $N = 5$ points.

The computation times can be found in Tables IV and VI, for the perfect and desired approximation cases, respectively. For both cases, as expected, exploiting the points' coplanarity (as discussed in Section III-B) leads to a great reduction in computation time. We can also notice that, as a rule of thumb, the symmetry of the chosen parameters plays a significant role in the resolution complexity, with more symmetric cases requiring much shorter computation times on average. Moreover, for the desired approximation only, we can see the exponential increase of time as we add new points to the configuration.

A. Outline of the simulations

The test cases presented below intend showing how the computations of the stable and unstable fixed points shine a new light on the dynamics of these systems. This is particularly true for systems that enforce the error to decrease, which is the case using the perfect approximation of the interaction matrix in the controller. For such error decreasing systems, fixed points are extrema of the error function, and saddles play a critical role: it is expected that saddles with index 1 (i.e., having 1 eigenvalue with positive real part and 5 eigenvalues with negative real part) have a stable manifold of dimension 5 that separates the regions of attraction of the global and local minima, see, e.g., [40]. The computation of these saddles allows gaining a global qualitative understanding of the dynamics. Although the exact location of the saddles stable manifolds is unknown[5], we can compute their unstable manifold since they are of dimension 1: it consists of two so-called heteroclinic trajectories[6], which exit the saddle in opposite directions toward the minima. Trajectories starting near the stable manifold of a saddle will first move toward the saddle, and then turn and exit toward a minimum following one of the two heteroclinic trajectories. This qualitative understanding of the dynamics is finer and more robust than just performing numerous simulations spanning a large portion of the state-space. We will therefore just simulate the heteroclinic orbits associated to the saddles: in the case of index 1 saddles, there is a pair of heteroclinic orbits; in the case of index 2 saddles, there is a two dimensional map of heteroclinic orbits and we will simulate a sample of them. The critical importance of saddles for error decreasing systems is also illustrated by performing statistics of trajectories starting near them to

[5]Fixed points and their index do not depend on the error decreasing controller, but the (un)stable manifolds actually depend on the controller. The computation of (un)stable manifolds of dimension greater than one is extremely difficult [41], and unfeasible for dimension five manifolds like here.

[6]In general, a heteroclinic trajectory (or orbit) is a trajectory that joins two different equilibrium points.

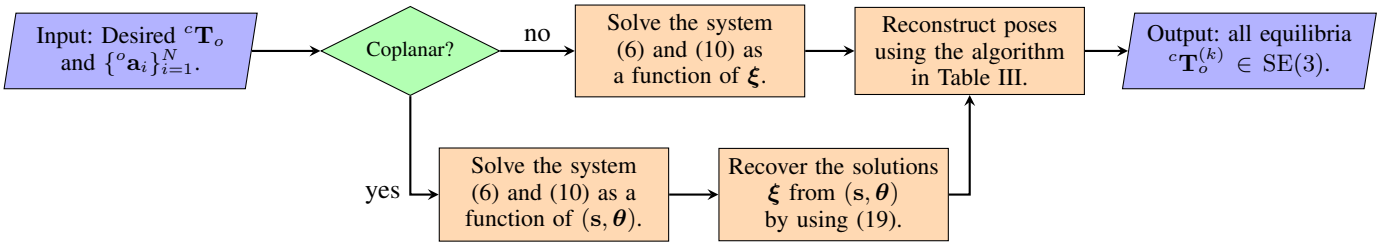


Fig. 1. Flowchart illustrating the equilibria computation's steps from start to finish.

TABLE IV
PERFECT APPROXIMATION TEST CASES CONSIDERED IN THE PAPER AND THEIR CORRESPONDING EQUILIBRIA'S COMPUTATION TIME, NUMBER AND STABILITY PROPERTIES.

	Configuration	Desired pose ID	Desired pose ${}^o\mathbf{t}_c$	Desired pose $\theta\mathbf{u}$	Comp. time	Equilibria result
1	 ${}^oP = \begin{bmatrix} 0.0 & 1.0 & 0.62 & 0.31 \\ 0.0 & 0.0 & 0.38 & 0.88 \\ 0.0 & 0.0 & 0.0 & 0.31 \end{bmatrix}$	$1a$	$\begin{bmatrix}-0.66, -0.43, 2.49\end{bmatrix}$	$\begin{bmatrix}-\pi, 0.0, 0.0\end{bmatrix}$	22.9h	2 stable, 2 saddles
		$1b$	$\begin{bmatrix}-3.96, 6.94, -0.41\end{bmatrix}$	$\begin{bmatrix}1.6, -0.4, 1.2\end{bmatrix}$	34.7h	2 stable, 2 saddles
2	 ${}^oP = \begin{bmatrix} 0.0 & 0.13 & 0.06 & 0.13 \\ 0.0 & 0.0 & 0.08 & -0.11 \\ 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$	$2a$	$\begin{bmatrix}-0.19, -0.01, -0.47\end{bmatrix}$	$\begin{bmatrix}-0.2, -0.2, 1.4\end{bmatrix}$	25.2h	2 stable, 2 saddles
		$2b$	$\begin{bmatrix}0.0, 0.0, -1.0\end{bmatrix}$	$\begin{bmatrix}0.0, 0.0, 0.0\end{bmatrix}$	2.7h	1 stable, 1 saddle
3	 ${}^oP = 0.07 \begin{bmatrix} -1 & 1 & 1 & -1 \\ k & k & -k & -k \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$3a$, $k = \frac{5}{7}$	$\begin{bmatrix}0.0, 0.0, -0.3\end{bmatrix}$	$\begin{bmatrix}0.0, 0.0, 0.0\end{bmatrix}$	58.5min	1 stable, 1 saddle
		$3b$, $k = \frac{6}{7}$	$\begin{bmatrix}0.0, 0.0, -0.3\end{bmatrix}$	$\begin{bmatrix}0.0, 0.0, 0.0\end{bmatrix}$	66.0min	1 stable, 1 saddle
		$3c$, $k = \frac{69}{70}$	$\begin{bmatrix}0.0, 0.0, -0.3\end{bmatrix}$	$\begin{bmatrix}0.0, 0.0, 0.0\end{bmatrix}$	103.4min	1 stable, 1 saddle
		$3d$, $k = 1$	$\begin{bmatrix}0.0, 0.0, -0.3\end{bmatrix}$	$\begin{bmatrix}0.0, 0.0, 0.0\end{bmatrix}$	13.1min	1 stable

confirm their equiprobability of converging toward different minima, see Table V. Finally, the video accompanying this paper illustrates the behavior of the system for each test case considered below.

Remark. *A complete qualitative understanding of the dynamics requires the verification that there exists no trajectory with camera position converging to infinity. This is granted with proper error function that have compact error level sets, but not granted for the image points investigated here, whose error function is not proper: a well-known counterexample is a set of coplanar points with parallel image plane desired configuration, and* 180 *degrees around the optical axis flipped initial condition whose trajectory moves backward to infinity [4].*

Section IV-B presents test cases using the perfect approximation interaction matrix in the control law. The three first test cases Case $1a$, Case $1b$ (four non-coplanar arbitrary points with two different desired poses) and Case $2a$ (four planar arbitrary points with non-parallel desired image plane), whose detailed descriptions are given in Table IV, all have one local minimum in addition to the global minimum and two saddles of index 1. They are represented in Figs. 2-4. These two index 1 saddles have stable manifolds of dimension 5 that form the boundary of the region of attractions of each minimum, as it can be guessed by looking at the corresponding 3D trajectories (Figs. 2a-4a). In each case, two heteroclinic orbits leave the saddles with almost equal initial images and converge toward different minima with their own but similar images, showing similar trajectories in the image (the similitude of the image views of the two heteroclinic orbits is really remarkable in Cases $1a$ and $2a$, as shown in Figs. 2b-2c and 4b-4c). However, the corresponding trajectories in $\mathrm{SE}(3)$ start in opposite directions and diverge from each other to eventually converge to different minima. Overall, the computation of the minima, saddles and heteroclinic orbits show the similar qualitative dynamics of these three example. Additionally, we numerically confirm our analysis by performing simulations with initial poses close to the saddles. For each of them, we sample 1000 initial poses uniformly in a small neighborhood

of the saddle[7] and verify how many converge to which stable equilibrium (see Table V for an overview).

Two cases of planar quadrilaterals with parallel desired image plane are then investigated: Case $2b$ (four arbitrary planar points) and Case 3 (a typical rectangle with centered desired pose). In both cases, these systems have only one minimum and one index 1 saddle, therefore showing a qualitative behavior very different than the previous test cases. Case 3 is investigated for different ratios for the rectangle sides, showing that the closer to a square the farther the saddle, see Figure 6. A closed-form formula (20) is given for the distance of this saddle with respect to the ratio, proving that it indeed diverges to infinity to finally disappear for a square in Case $3d$.

Section IV-C presents test cases using the desired pose interaction matrix in the control law (see Table VI). This system is not anymore energy decreasing, hence in spite of a simpler expression due to the usage of a fixed interaction matrix, it has more complex dynamics. Surprisingly, all tested non-planar cases (four and five points non-planar configurations, with two desired poses in each case) have exactly one attractive fixed point and one saddle. Even more surprisingly, the saddle now has index 2, meaning that a surface of heteroclinic trajectories leaves the saddle, here all converging toward the desired pose. All of them have arbitrarily similar initial images and identical final images, but run very different trajectories both in $\mathrm{SE}(3)$ and in image space, see Fig. 7a and Figs. 8a-8b, as well as the accompanying video. Together with some trajectories that do not converge due to physical constraints, this quantitative analysis shows a complex global dynamics of this controller.

B. Perfect approximation test cases

1) Test case 1*, non-planar configuration:* In this example, we consider the case where the 4 points are arranged in an arbitrary non-planar configuration. Two different desired poses are considered, one which is such that the optical axis of the camera is perpendicular to one of solid's faces (Case $1a$ in Table IV) while the other is completely arbitrary (Case $1b$). In both cases, we have 4 equilibria, one of which coincides with the desired pose (which is stable), a second stable equilibrium (a local minimum) and two unstable equilibria (specifically, two saddles).

It is possible to see the computed equilibria and the corresponding heteroclinic orbits for Cases $1a$ and $1b$ in Figs. 2-3. We can see (Figs. 2a-3a) that all trajectories converge to a stable equilibrium, and, as expected, roughly half of them converge to the desired pose (see Table V). As for the trajectories produced in the image, we can note in Figs. 2b-2c and Figs. 3b-3c that the controller produces as expected straight line trajectories for the four image points for converging either to the global minimum or to the local one.

Interestingly, comparing the global and local minima in image space, we can see that, as it might be expected, they are very similar for Case $1a$ (Fig. 2d), while they are surprisingly different for Case $1b$ (Fig. 3d).

[7]We consider a maximum displacement of $1\,\mathrm{cm}$ in position and $0.5°$ in orientation.

TABLE V
PERCENTAGE OF TRAJECTORIES (OUT OF 1000 SIMULATIONS) STARTING IN A NEIGHBORHOOD OF THE SADDLES AND CONVERGING TO THE DESIRED POSE; THE REMAINING ONES CONVERGE TO THE LOCAL MINIMUM WHEN IT EXISTS.

	Ex. $1a$	Ex. $1b$	Ex. $2a$	Ex. $2b$
Sad. 1	51.6%	49.6%	51.3%	100%
Sad. 2	49%	51.4%	49.8%	N.A.

2) Test case 2*, planar quadrilateral:* In this example, we consider the case where the 4 points are coplanar (but non-aligned) and arranged in an arbitrary quadrilateral configuration. Once again, two different desired poses are considered, one which is completely arbitrary (Case $2a$ in Table IV) and the other where the quadrilateral is parallel to the image plane (Case $2b$). In the first case, we have again 4 equilibria, two of which are stable (being one of them the desired pose), while the other two are saddles. However, in the second case, we only have 2 equilibria, one of which coincides with the desired pose (which is stable), while the other is a saddle. For the first case, we can note the symmetry between the desired pose and the local minimum with respect to the object plane (even if it is not perfect) and the fact that the saddles are almost parallel to the object plane. These results are fully coherent with the state-of-the art recalled in Section I that a local minimum exists for a planar object in the non-parallel case. In this specific example, we show for the first time that the local minimum is unique and that it does not exist in the parallel case. We conjecture that this might be a general property for this class of cases, even though it still remains an open question.

The computed equilibria for Cases $2a$ and $2b$ are illustrated in Figs. 4-5, together with trajectories starting in a neighborhood of the saddles. The overall behavior in Case $2a$ matches the ones in Cases $1a$ and $1b$. In Case $2b$, however, we only have one saddle and one stable equilibrium, and all the trajectories converge to the desired pose; interestingly, though, they still do so following two distinct orbits. In Case $2a$, we can note that the positions of the four points in the image are almost the same for the two stable equilibria and for the saddle located at their middle, while they are very different for the other saddle (as well as for the single saddle in Case $2b$). Finally, the percentage of simulated trajectories that converge to the desired pose are given in Table V, where we can see that, as expected, roughly half of the trajectories converge to the local minimum, when it is present. Note that similar results are obtained when the 4 points are arranged in a rectangle or square configuration, with the exception of the particular case exhibited below.

Finally, for Case $2a$, the behavior of the system around the saddle S_2 has also been validated experimentally on a 6 DoF Gantry robot with an eye-in-hand INTEL(R) REALSENSE D435 camera, giving equivalent results with respect to the simulations. Notably, while the simulations used two close initial poses, in the experiment we used the very same initial pose, with the intrinsic imprecision of the physical system

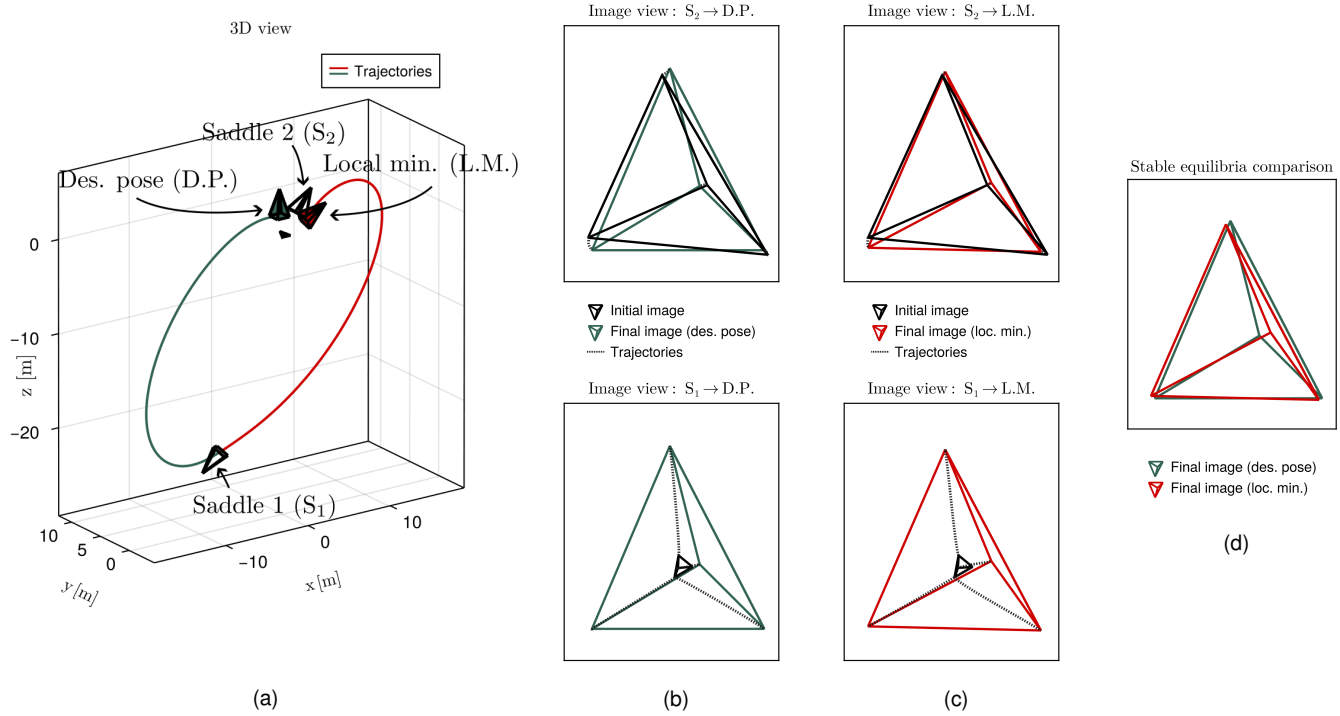


Fig. 2. Computed equilibria and simulations around the saddles for Case 1a. (a): 3D view, computed equilibria and heteroclinic orbits for the pseudo-inverse controller. (b) and (c): image view, trajectories converging to the desired pose (b) and local minimum (c). (d): comparison between local minimum and desired pose in image space.

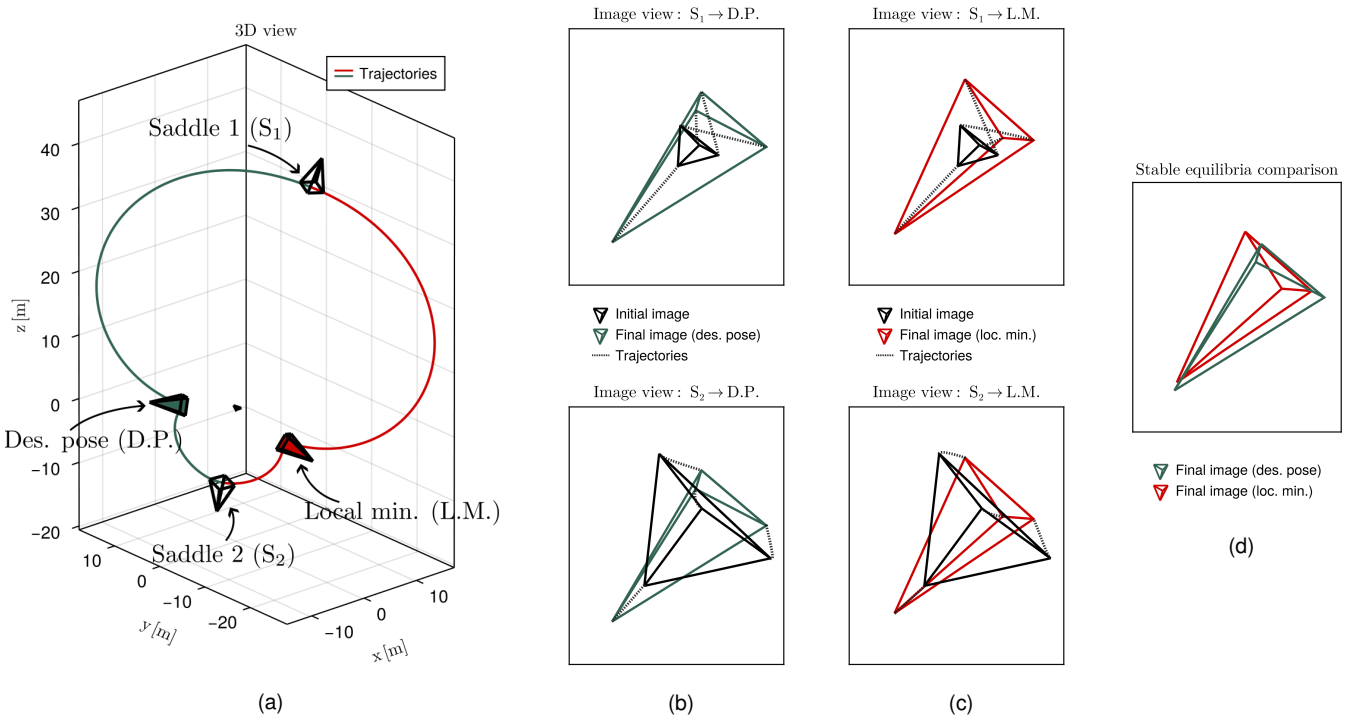


Fig. 3. Computed equilibria and simulations around the saddles for Case 1b. (a): 3D view, computed equilibria and heteroclinic orbits for the pseudo-inverse controller. (b) and (c): image view, trajectories converging to the desired pose (b) and local minimum (c). (d): comparison between local minimum and desired pose in image space.

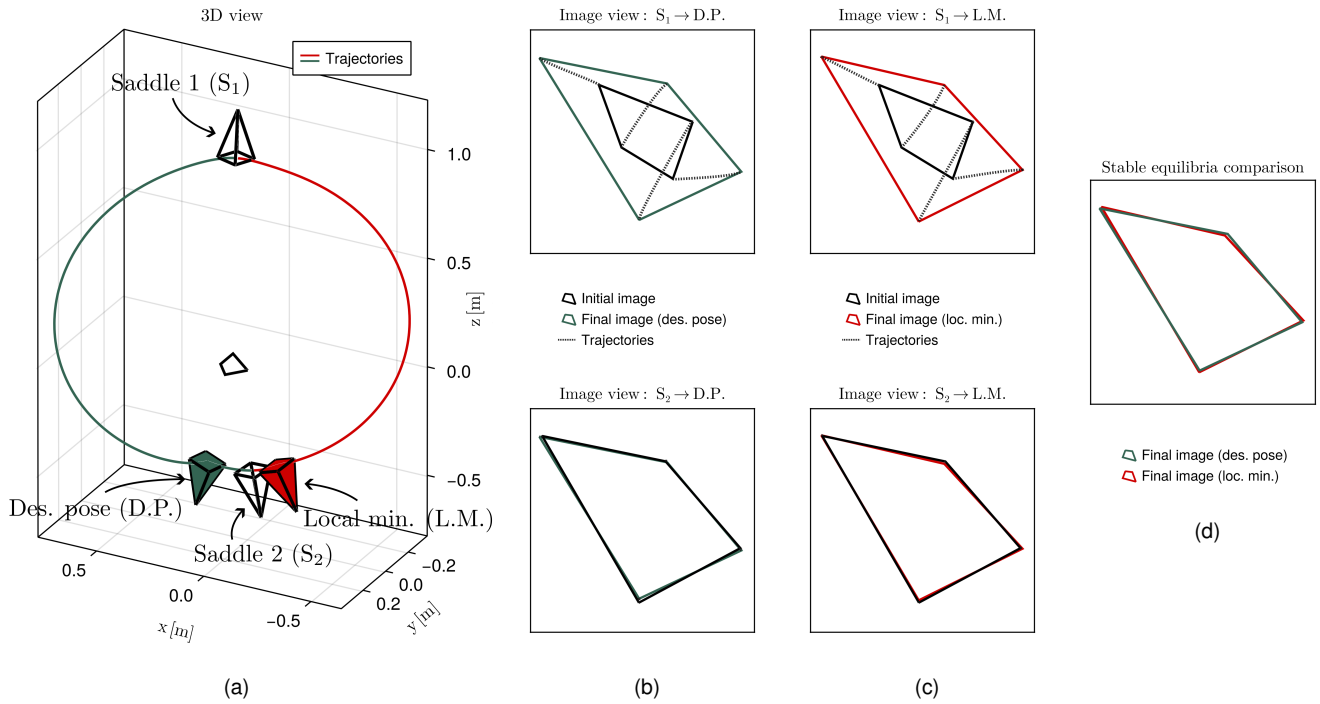


Fig. 4. Computed equilibria and simulations around the saddles for Case $2a$. (a): 3D view, computed equilibria and heteroclinic orbits for the pseudo-inverse controller. (b) and (c): image view, trajectories converging to the desired pose (b) and local minimum (c). (d): comparison between local minimum and desired pose in image space.

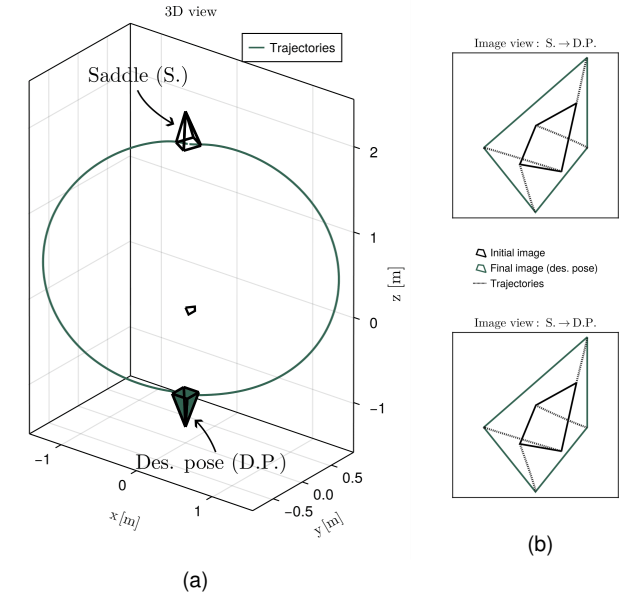


Fig. 5. Computed equilibria and simulations around the saddles for Case $2b$. (a): 3D view, computed equilibria and heteroclinic orbits for the pseudo-inverse controller. (b): image view, trajectories converging to the desired pose.

causing random convergence to either stable equilibrium. The interested reader may refer to the end of the accompanying video to observe the experiment.

3) Test case 3*:* In this last example, we want to showcase a peculiar behavior obtained while studying the classical case where the 4 points are coplanar and arranged in a rectangle, and where the desired pose is such that the rectangle is centered and parallel to the image plane. We fix the desired pose and we consider four different points configurations, starting with a narrow rectangle (Case $3a$ in Table IV) and making it gradually wider (Cases $3b$ and $3c$) until it becomes a square (Case $3d$). For the tested rectangles, there always exist 2 equilibria, one of which coincides with the desired pose (which is stable), while the other is a saddle. As it can be seen in Fig. 6, for all three cases the unstable equilibrium is a "mirror image" of the desired pose, i.e., it is rotated $180°$around the camera x-axis and it is such that the optical axis is aligned with the desired pose's optical axis. Furthermore, the trajectories of the four points in the image are no more pure straight line but slightly differ starting from one side of the saddle or the other. We can also note that the four points become aligned in the image when the camera optical center crosses the rectangle plane.

Interestingly, as the rectangle gets closer to a square, the saddle gets pushed farther and farther and, when it becomes a square (Case $3d$), we have only one equilibrium, which coincides with the desired pose (which, again, is stable by design). Although, to the best of our knowledge, this is the first time that the uniqueness of the equilibrium is rigorously confirmed for the classical parallel and centered square configuration, it is important to note that this does not imply that the system is globally asymptotically stable; as already said, there exist unstable trajectories for this system, typically for a rotation of $180°$ around the optical axis [4].

The previous result suggests that the saddle is moved back-

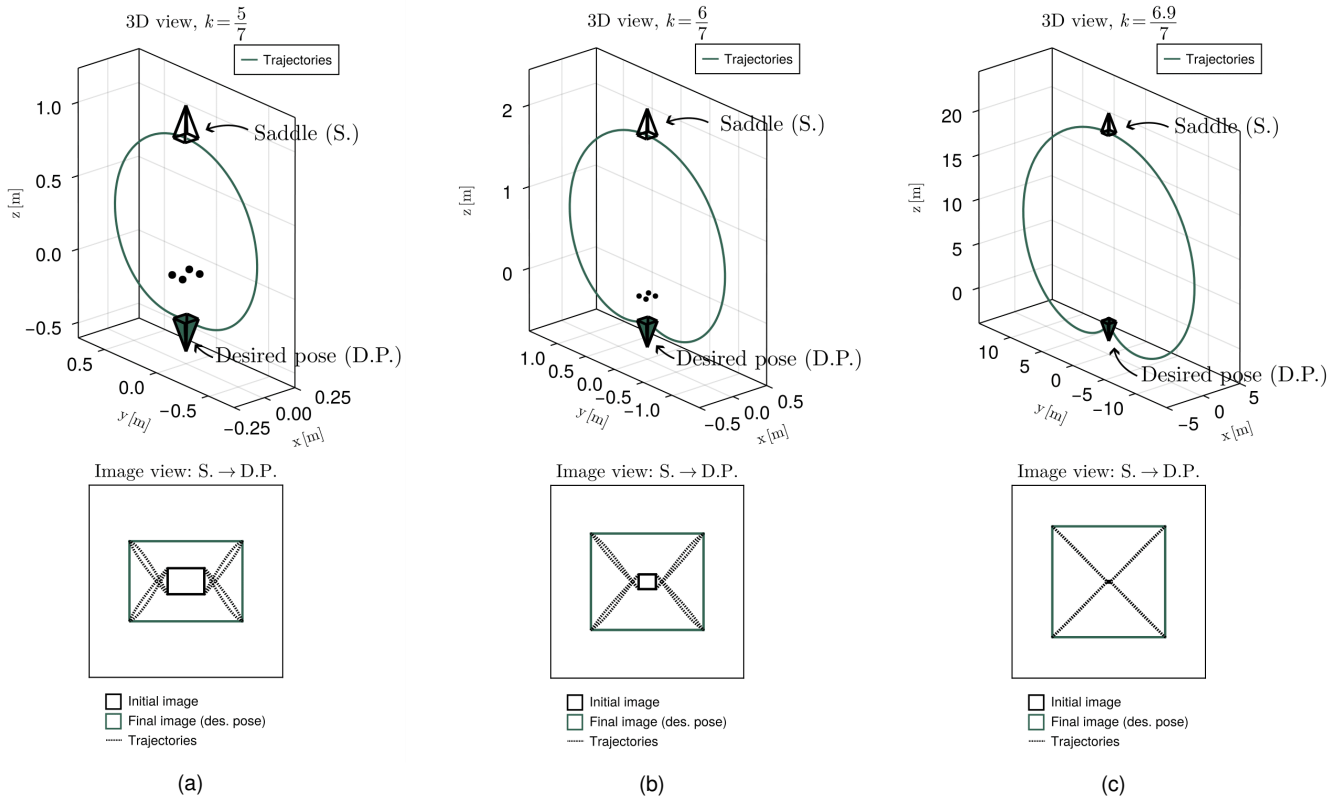


Fig. 6. Cases $3a$-$3c$ (left to right). On the top row: 3D view, computed equilibria and heteroclinic orbits for the pseudo-inverse controller. On the bottom row: image view, trajectories converging to the desired pose.

ward at infinity for the square configuration, which motivated us to investigate more deeply this particular case of a parallel centered square. We consider the ratio k between the two sides of the rectangle as a parameter, and we restrict the camera/square poses to a pure translation along the optical axis while fixing their orientation to the saddle's one. Doing so, the extended features $\boldsymbol{\xi}$ only depend on k and the depth Z w.r.t. the points plane, which significantly simplifies the system (6) and (10). It is possible to show that this new system is equal to zero if and only if:

$$Z = Z_d \frac{1+k^2}{|1-k^2|}, \tag{20}$$

where $Z_d > 0$ represents the desired depth w.r.t. the points plane. This means that there is exactly one equilibrium that lies on this restriction for any choice of $k \neq 1$, and it proves that the equilibrium's position moves backwards and tends to infinity as the points get closer and closer to a square configuration. On the other hand, it does not provide any information on the presence of other equilibria lying outside of this restriction, even though we conjecture that no other fixed points exist from the four examples we have considered in this case.

C. Desired pose approximation test cases

Similarly to the previous section, we study three examples of points' configurations, analyzing a representative set of desired poses for each example. Table VI summarizes the examples' parameters. Cases $1a$ and $1b$ are taken directly from the previous section. Cases $2a$ and $2b$ consider the same desired poses as Case 1 for a 5 points' configuration, obtained by adding one point to the previous 4 points. Finally, Case 3 showcases a peculiar behavior that arises for this class of controllers.

1) Test case 1*,* 4 *points non-planar configuration:* In this example, as already said, we reanalyze Case 1 from Section IV-B, where 4 points are arranged in an arbitrary non-planar configuration. The same desired poses are considered, one which is such that the optical axis of the camera is perpendicular to one of solid's faces (Case $1a$ in Table VI) while the other is completely arbitrary (Case $1b$). Contrarily to the perfect approximation controller, we now have 2 equilibria for both cases, one of which coincides with the desired pose (which is stable) and one unstable equilibrium (specifically, a saddle). These equilibria are illustrated in Figs. 7a-7b, where we also show a selection of trajectories, both in camera state- and image space, starting in a close proximity of the saddle. Comparing these trajectories with the ones obtained with the perfect approximation controller, we have a striking qualitative difference between the trajectories' behavior in the two settings. For both desired poses, the trajectories seem to develop on a surface connecting the saddle to the desired pose (even though the surfaces appear to be significantly more regular for Case $1b$ than Case $1a$). This behavior is linked to the dimension of the saddle's unstable manifold: checking the eigenvalues of the system's Jacobian computed at the saddle,

TABLE VI
DESIRED POSE APPROXIMATION TEST CASES CONSIDERED IN THE PAPER AND THEIR CORRESPONDING EQUILIBRIA'S COMPUTATION TIME, NUMBER AND STABILITY PROPERTIES.

	Configuration		Desired pose			Comp. time	Equilibria results
			ID	${}^o\mathbf{t}_c$	$\theta\mathbf{u}$		
1		${}^oP = \begin{bmatrix} 0.0 & 1.0 & 0.62 & 0.31 \\ 0.0 & 0.0 & 0.38 & 0.88 \\ 0.0 & 0.0 & 0.0 & 0.31 \end{bmatrix}$	$1a$	$\begin{bmatrix}-0.66, -0.43, 2.49\end{bmatrix}$	$\begin{bmatrix}-\pi, 0.0, 0.0\end{bmatrix}$	27.3s	1 stable, 1 saddle
			$1b$	$\begin{bmatrix}-3.96, 6.94, -0.41\end{bmatrix}$	$\begin{bmatrix}1.6, -0.4, 1.2\end{bmatrix}$	116s	1 stable, 1 saddle
2		${}^oP = \begin{bmatrix} 0.0 & 1.0 & 0.62 & 0.31 & -0.31 \\ 0.0 & 0.0 & 0.38 & 0.88 & 0.5 \\ 0.0 & 0.0 & 0.0 & 0.31 & -0.31 \end{bmatrix}$	$2a$	$\begin{bmatrix}-0.66, -0.43, 2.49\end{bmatrix}$	$\begin{bmatrix}-\pi, 0.0, 0.0\end{bmatrix}$	254.8s	1 stable, 1 saddle
			$2b$	$\begin{bmatrix}-3.96, 6.94, -0.41\end{bmatrix}$	$\begin{bmatrix}1.6, -0.4, 1.2\end{bmatrix}$	2.2h	1 stable, 1 saddle
3		${}^oP = \begin{bmatrix} -0.07 & 0.07 & 0.07 & -0.07 \\ 0.07 & 0.07 & -0.07 & -0.07 \\ 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$	$3a$	$\begin{bmatrix}0.0, 0.0, -0.5\end{bmatrix}$	$\begin{bmatrix}0.0, 0.0, 0.0\end{bmatrix}$	0.07s	1 stable, 4 saddles
			$3b$	$\begin{bmatrix}0.0, 0.0, -0.3\end{bmatrix}$	$\begin{bmatrix}0.0, 0.0, 0.0\end{bmatrix}$	0.07s	1 stable, 4 saddles
			$3c$	$\begin{bmatrix}0.0, 0.0, -0.14\sqrt{2}\end{bmatrix}$	$\begin{bmatrix}0.0, 0.0, 0.0\end{bmatrix}$	0.08s	1 stable, 4 saddles
			$3d$	$\begin{bmatrix}0.0, 0.0, -0.1\end{bmatrix}$	$\begin{bmatrix}0.0, 0.0, 0.0\end{bmatrix}$	0.07s	1 stable

we can see that in both cases we have two positive eigenvalues.

Moreover, with respect to the perfect approximation controller, we have a new behavior that arises, which is the presence of *failed* trajectories, i.e., trajectories that come too close to the object so that the depth Z of at least one object point becomes null (shown with dashed lines in the figures). This phenomenon, however, is not surprising: the employed control law is based on a constant and very coarse approximation of the true interaction matrix, and, by design, we expect it to be effective only when the camera is reasonably close to the desired pose. In general, we have no guarantees for trajectories with arbitrary initial conditions, and, as we will see, this behavior arises in almost all the examples considered in this section. Additionally, considering the trajectories in the image, we see that the points do not move at all along straight lines (while it is the case for the perfect controller). Once again, this validates the approximated nature of this class of controllers, which leads to more unpredictable (but, also, richer) behavior.

Similar examples where the 4 points are arranged in a coplanar rectangle configuration lead to comparable results with respect to Case 1, the reason why these cases were left out from the current discussion. The interested reader can find one such example in the accompanying video.

2) Example 2*,* 5 *points non-planar configuration:* In this example, we study a variation of Cases $1a$ and $1b$. We consider a configuration of 5 points, formed by the 4 points considered in the previous example and an additional, arbitrary point, and we use the very same desired poses (Cases $2a$ and $2b$ in Table VI). Again, in both cases, we have 2 equilibria, a stable one (coinciding with the desired pose) and a saddle. Even though we have significant similarities with the previous example, quite surprisingly the saddles do not bear any evident resemblance to what we found above.

In Figs. 8a-8b, we can see the computed equilibria and a selection of trajectories for Cases $2a$ and $2b$. While in Case $2b$ the trajectories seem to develop on a surface connecting the saddle to the desired pose, matching the behavior of Cases $1a$ and $1b$, in Case $2a$ the trajectories quickly converge to one single orbit, as it was the case for the saddles seen in Section IV-B. Once again, this behavior is due to the dimension of the saddle's unstable manifold: in Case $2a$ we have an index 1 saddle, while there are two positive eigenvalues in Case $2b$.

Finally, we can see that the trajectories in the image are far from developing along straight lines, and they present an erratic behavior that changes conspicuously depending on the initial pose, which showcases, once again, the approximated nature of this control law.

3) Example 3*:* In this last example, we want again to showcase a curious behavior that arises in the classical case where the 4 points are coplanar and arranged in a square, and where the desired pose is such that the square is centered and parallel to the image plane. We consider four different desired poses, all with the same orientation and lying on the same axis, with only the distance from the points plane changing: we start farther away (Case $3a$ in Table VI) and we get increasingly closer (Cases $3b$ to $3d$) to the points plane. In the first three cases, we have 5 equilibria, four of which are saddles, while the fifth one is the desired pose (stable by design). All the saddles lie on the points plane, and they are arranged symmetrically with respect to the points. In the fourth case, instead, we only have 1 equilibrium, coinciding with the desired pose. By looking at Fig. 9, we can see that, as the

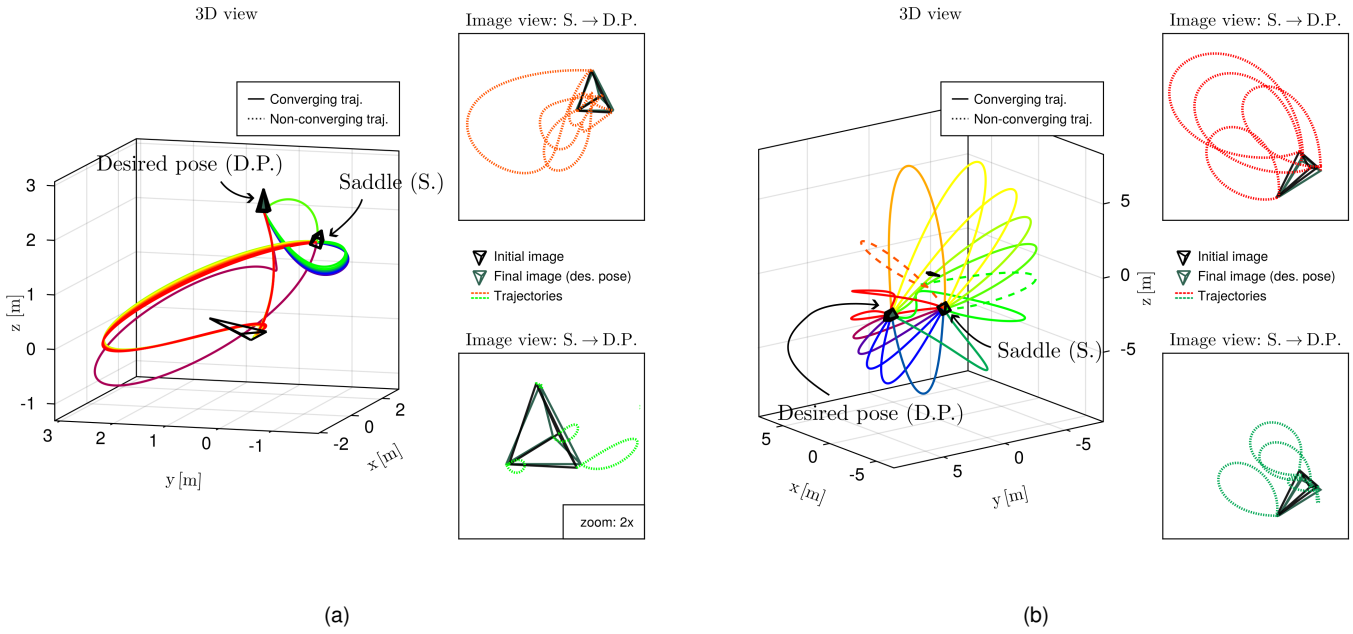


Fig. 7. Computed equilibria and a selection of heteroclinic orbits, both in 3D and image space. (a): Case 1a. (b): Case 1b.

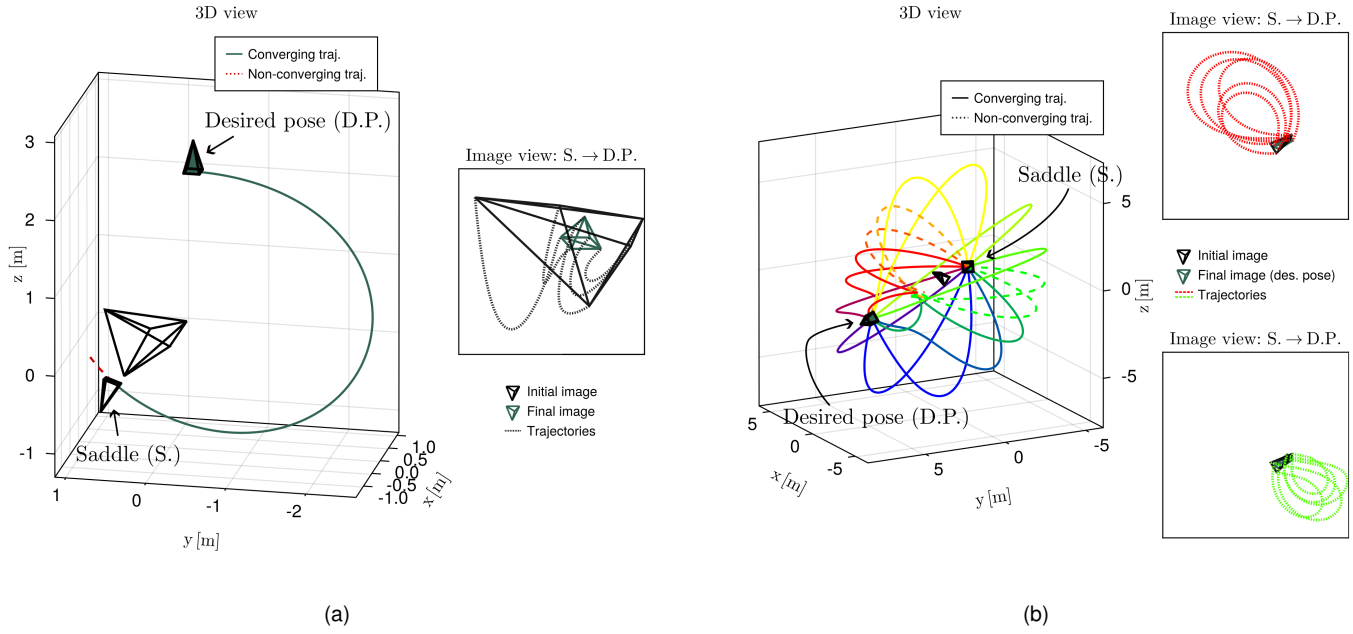


Fig. 8. Computed equilibria and a selection of heteroclinic orbits, both in 3D and image space. (a): Case 2a. (b): Case 2b.

desired pose approaches the points plane, the four saddles get closer to the four points (Cases 3a and 3b), until they reach a position where they are superposed to the points (Case 3c). Drawing the desired pose even closer, the saddles eventually disappear (Case 3d), or, more precisely, they become such that at least one point is behind the camera optical center, which excludes them to be a feasible solution.

In all cases, the four saddles share the same relative orientation, and only their distance to the points changes from case to case, which motivated us to investigate this example more deeply. In the same spirit of the analysis carried out in Section IV-B3, we consider the desired depth Z_d as a parameter, and we restrict the camera's position to be on one of the axes where the saddles' lie[8] and we fix its orientation to the saddle's one. With this restriction, the extended features $\boldsymbol{\xi}$ depend only on the (signed) distance d w.r.t. the closest point, with $d > 0$ when the point lies in front of the camera (i.e., when the point's depth in camera-frame is positive).

[8]We can choose any of the four axes; it has no impact on the result due to the symmetry of the problem.

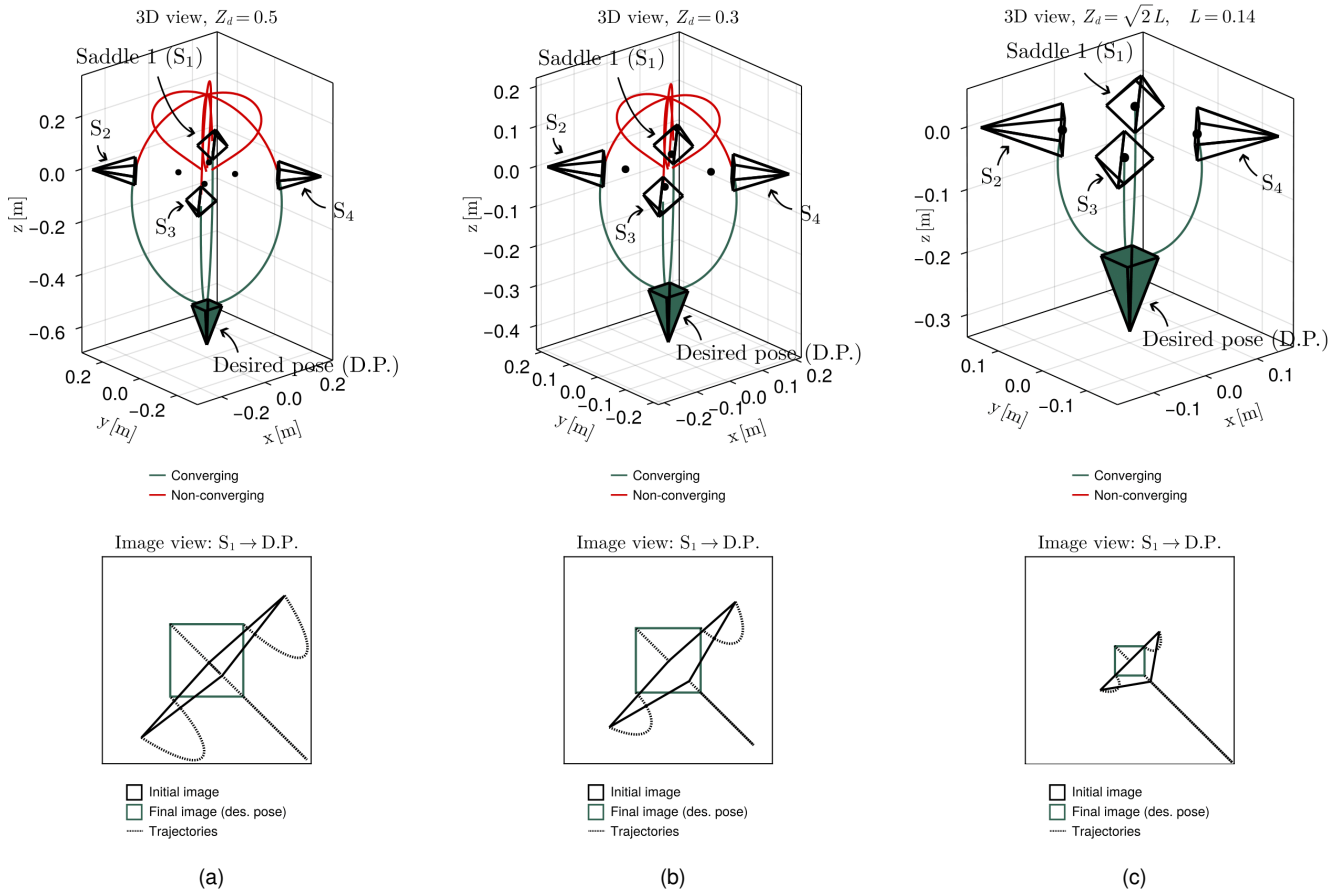


Fig. 9. Cases 3a-3c (left to right). On the top row: 3D view, computed equilibria and heteroclinic orbits for the pseudo-inverse controller. On the bottom row: image view, one of the heteroclinic orbits.

Substituting into the system (6) and (10), it is possible to show that the equilibrium condition is equal to zero if and only if:

$$d = \frac{1}{2}\left(Z_d - \sqrt{2}L\right), \tag{21}$$

where $L > 0$ is the length of the square's side. This means that these four, symmetrical saddles are present for every choice of $Z_d \geq \sqrt{2}L$, and, as Z_d decreases, the saddles positions get closer and closer to the points, until they eventually become superposed to them at the equality. Interestingly, these saddles are solutions of (6) and (10) even in the case where $d < 0$, and they tend to converge to the center of the square as Z_d tends to 0; however, these solutions lose physical meaning, since one of the points is behind the camera.

V. Sensitivity Analysis

In this section, we analyze the impact of perturbations on the study of equilibria we have developed so far. In case of uncertainties only on the point depths Z_i, it is clear that the desired pose is a robust equilibrium. Indeed, by looking at Eq. (4), $\mathbf{e} = \mathbf{0}$ is sufficient to have an equilibrium, regardless of $\mathbf{Z}$. By the system's continuity, we have that, for acceptable levels of perturbation, the system still precisely converges to the desired pose. However, for all other equilibria and/or in case of image noise, we expect that perturbation prevents the system from converging exactly to the equilibrium, since the presence of noise disrupts the deterministic structure of the system. However, if we have reasonable bounds on the noise amplitude, it is possible to show that stable equilibria become *attractors*, i.e., invariant and locally attractive sets for the system [42], meaning that we still expect it to converge towards a neighborhood of the stable equilibria.

To verify this hypothesis and quantify the impact of uncertainties, a series of simulations afflicted by noise has been conducted, distinguishing between errors only in depth estimates and errors in all extended features. In particular, we considered additive uniform noise with zero mean and realistic amplitudes, as detailed in the top rows of Table VII. The sensor noise levels correspond to an error of ± 2 and ± 5 pixels, respectively, for a camera with 700 as focal length-pixel size ratio. Additionally, we only present the results for Case 1a (see Table IV), as we have observed that the results are similar for all examples.

In order to study the behavior around the stable equilibria, we firstly perform a series of simulations starting from poses randomly chosen within a large neighborhood of the two saddle points. In particular, we sample 1000 random initial poses for each saddle point, with a maximum displacement of 1 m in position and 15° in orientation. As expected, all simulations converge towards one of the stable equilibria, ultimately

TABLE VII
OVERVIEW OF THE ANALYSIS OF CASE $1a$ UNDER PERTURBATION, DEPENDING ON THE TYPE AND LEVEL OF NOISE.

	$\mathbf{Z}$ only	$(\mathbf{s}, \mathbf{Z})$ simultaneously	
Noise range on s_i	0	± 0.003	± 0.007
Noise range on Z_i	± 5cm	± 5cm	± 5cm
Dist. to des. pose	0cm	0.09cm	0.33cm
Ellipsoid vol.	0cm^3	37.47cm^3	525.5cm^3
Dist. to loc. min.	0.44cm	0.41cm	0.21cm
Ellipsoid vol.	1.22cm^3	21.49cm^3	130.7cm^3

getting trapped inside a neighborhood of the equilibrium. Out of the initial poses sampled around the saddle point S_1, 57.1% converge towards the desired pose and 42.9% towards the local minimum, while the convergence percentage for the poses sampled around S_2 stands at 41% and 59% respectively, considering the highest noise level. Then, we assess these neighborhoods' *centers* (representing the average distance to the equilibrium) and *volumes* (representing the trajectories' spread), by running simulations for approximately 100 times the unperturbed convergence time and identifying the smallest ellipsoid containing all points after an initial transient. The values for the average distance to the equilibrium (centers) and volumes are listed in Table VII for each equilibrium and choice of perturbation. Interestingly, we notice that the ellipsoids' radii seem to increase linearly with respect to the noise level on the visual features. Additionally, we show an illustrative selection of simulations in Fig. 10, for the case of noise on $\mathbf{Z}$ alone and simultaneously on $\mathbf{s}$ and $\mathbf{Z}$, for the highest level of image noise. Our intuition of the perturbations' impact is confirmed, since the qualitative behavior of the perturbed system matches the unperturbed case, while the deviation of the steady-state poses with respect to the unperturbed equilibria is low and even null for the global minimum when only depth errors are introduced in the controller.

Furthermore, in order to showcase the importance of saddle points even in the perturbed case, we want to assess whether they maintain their role as separators of the regions of attraction. We thus selected a small grid of initial poses centered on the saddle point S_2, and for each of these points, we ran 1000 simulations with that point as the initial pose. We then counted, for each point, how many simulations converge within a neighborhood of the desired pose and how many are attracted to the local minimum, allowing us to construct an estimation of the probability density of the convergence towards the desired pose. A comparison between the unperturbed and perturbed cases can be seen in Fig. 11. As expected, in the unperturbed case, there is a deterministic separation between the regions of attraction of the two stable equilibria, with the boundary between them represented by the stable manifold of the saddle point. In the perturbed case, we have a qualitatively similar behavior, with the difference that the boundary between the two regions is no longer a sharp surface but a rather "blurred" region, where the convergence to either equilibrium becomes stochastic, depending on the perturbation itself. Finally, we highlight that the same behavior can be observed in practice, as discussed at the end of Sec. IV-B2 and as it can be seen at the end of the accompanying video.

VI. CONCLUSION

In this paper, we proposed a strategy to systematically assess all the equilibria, both stable and unstable, of image point-based visual servoing systems, for a class of 6 well-known controllers employed in the literature. Our strategy relies on defining the equilibrium condition directly in the extended features-space $\boldsymbol{\xi} = (\mathbf{s}, \mathbf{Z})$, to which we added the *feasibility constraints* that let us solve the system of equations directly on the subset of feasible extended features. By recasting the system of equations in polynomial form, we were able to employ a state-of-the-art certified polynomial solver, based on Gröbner basis theory, to find all the solutions of the system in a rigorous way for a representative class of test cases. Even though this modelization offers a great degree of simplification, the computational complexity of the equilibria computation currently restricts us to 4 points for the perfect approximation controller and 5 points for the desired pose approximation controller.

Nevertheless, our strategy grants us the ability to analyze a large class of classical situations in a more profound way, shining a new light on the behavior of these systems. In the authors knowledge, this is the first time that all stable and unstable equilibria of these systems have been identified, which allows us to provide a deeper qualitative understanding on the system's dynamics, in particular for the perfect approximation case. Moreover, it opens the door to a rich set of previously unknown behaviors, as well as providing a new way to study known phenomena in the field.

For instance, by formally assessing the uniqueness of the local minimum (Case $2a$ in Section IV-B) and the absence of it (Case $2b$ in Section IV-B) for a specific case of a planar quadrilateral configuration, we provide new evidence to the long-standing conjecture that this is always the case for planar objects. In particular, the absence of local minima in the parallel configuration is a promising result that gives hope to demonstrate formally the stability of the system in a very large workspace. Additionally, thanks to our new formulation of the equilibrium condition, we were able to explicitly compute the poses of some saddles in a parametric way, in particular cases of planar rectangular and square configurations that depend on a parameter (Case 3 in Section IV-B and Case 3 in Section IV-C). Finally, the identification of the unstable equilibria for the desired pose approximation controllers allowed us to explore the dynamical richness of these systems, shining light on what happens beyond the usual bounds of the desired pose's proximity, where the approximation introduced by these controllers is expected to be modest.

Hopefully, this might be just the beginning of a line of research that aims to build up a global understanding of

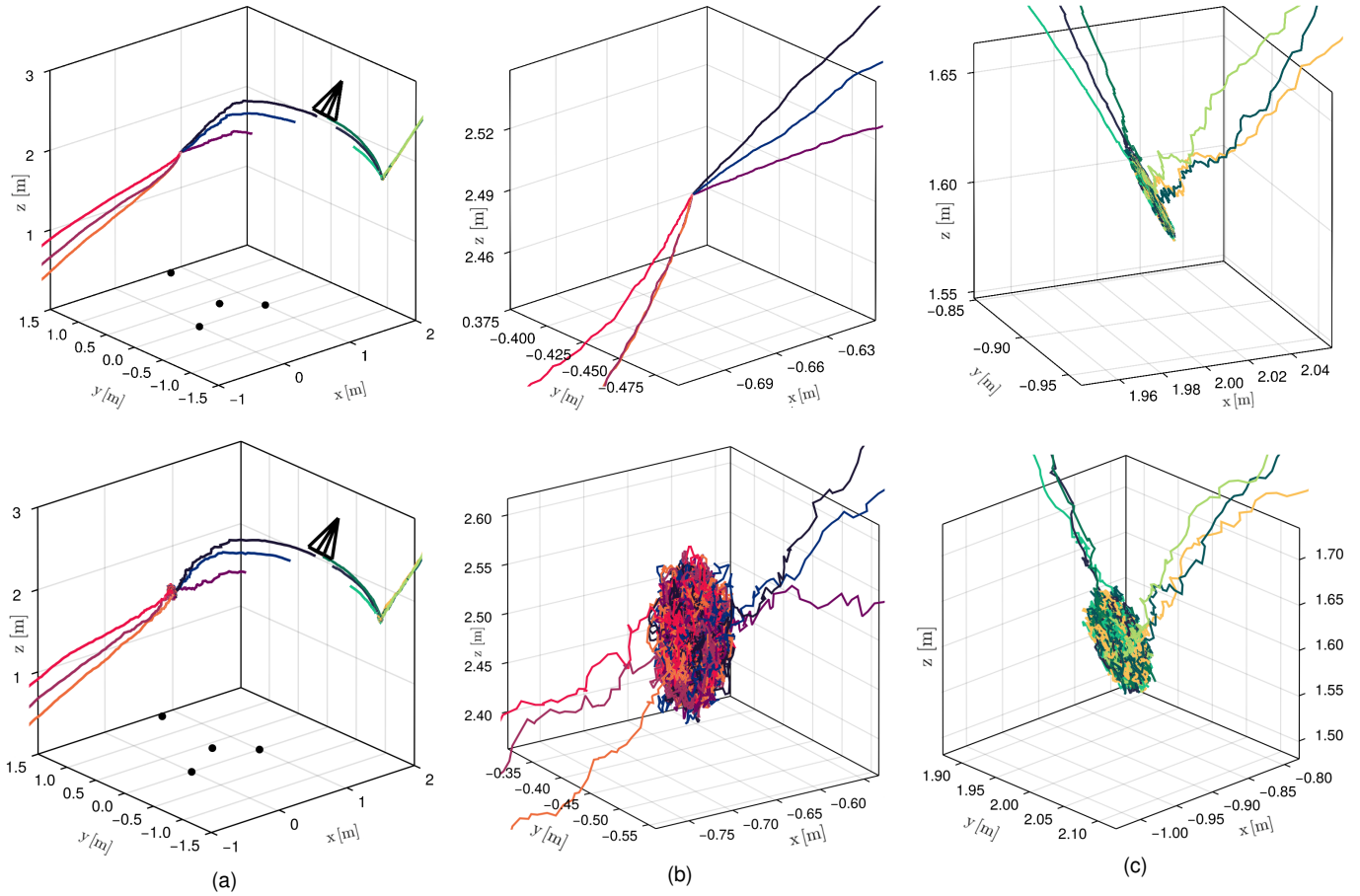


Fig. 10. Selection of 12 simulations with noise on $\mathbf{Z}$ only on top row and with simultaneous noise on $\mathbf{s}$ and $\mathbf{Z}$ on bottom row. (a): Trajectories overview, where 6 of the initial poses were chosen around each saddle point. (b)-(c): Zoom-in at the desired pose and local minimum, respectively.

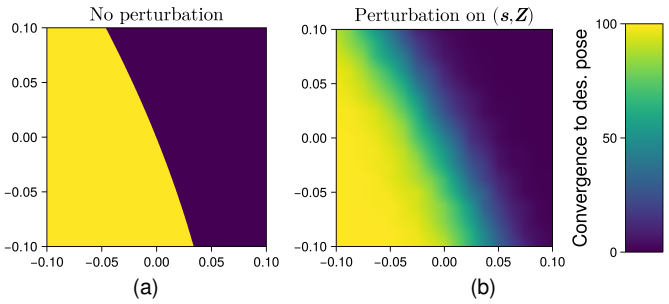


Fig. 11. Probability density of the convergence towards the desired pose for a 2-dimensional neighborhood centered on the saddle S_2, Case $1a$. (a): No noise. (b): Simultaneous noise on $\mathbf{s}$ and $\mathbf{Z}$.

the dynamics of these systems. As thoroughly discussed in Section IV-A, for the controllers that use the true interaction matrix, localizing the saddles allows us to start studying the stable equilibria's regions of attraction, since the stable manifold of the saddles represents the boundary between these regions. The characterization of these regions could prove itself to be crucial in the design of controllers that can escape the local minima, since it would allow the system to detect whether it is inside the local minimum's region of attraction. On the other hand, finding explicit conditions that guarantee the stable equilibrium's uniqueness could pave the way for the design of tasks for which the system always converges to the desired pose. In the same spirit, the desired pose approximation controllers' examples considered in Section IV-C make us hope that there is only one stable equilibrium in these cases. If this were to be confirmed, gaining a better understanding on the failed trajectories would become crucial in order to discuss the stability properties of these systems.

From a computational point of view, we believe that, as the solving algorithms and hardware evolve, attacking examples with configurations of 5 points for the perfect approximation controller will be within reach in the near future. Another interesting, but currently too computationally heavy, problem is to clear up how the equilibria move as we let one or more parameters of the problem change (for instance, by finding a parametric formulation of the equilibria w.r.t. the size or shape of the object).

REFERENCES

[1] F. Chaumette and S. Hutchinson. Visual servo control, part I: Basic approaches. *IEEE Robotics & Automation Magazine*, 13(4):82–90, Dec. 2006.

[2] W. Wilson, C. Hulls, and G. Bell. Relative end-effector control using Cartesian position based visual servoing. *IEEE Trans. on Robotics and Automation*, 12(5):684–696, Oct. 1996.

[3] B. Thuilot, P. Martinet, L. Cordesses, and J. Gallice. Position based visual servoing: Keeping the object in the field of vision. In *IEEE Int. Conf. on Robotics and Automation, ICRA'02*, pages 1624–1629, Washington DC, May 2002.

[4] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In *The Confluence of Vision and Control*, pages 66–78. LNCIS Series, No 237, Springer-Verlag, 1998.

[5] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 D visual servoing. *IEEE Trans. on Robotics and Automation*, 15(2):238–250, Apr. 1999.

[6] F. Chaumette and E. Malis. 2 1/2 D visual servoing: a possible solution to improve image-based and position-based visual servoings. In *IEEE Int. Conf. on Robotics and Automation, ICRA'00*, pages 630–635, San Francisco, CA, Apr. 2000.

[7] V. Kyrki, D. Kragic, and H. Christensen. New shortest-path approaches to visual servoing. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'04)*, pages 349–354, Sendai, Japan, Sep. 2004.

[8] P. Martinet, J. Gallice, and D Khadraoui. Vision based control law using 3D visual features. In *World Automation Congress*, pages 497–502, Montpellier, France, May 1996.

[9] E. Cervera, A.P. del Pobil, F. Berry, and P. Martinet. Improving image-based visual servoing with three-dimensional features. *The Int. Journal of Robotics Research*, 22(10-11):821–839, 2003.

[10] F. Schramm, G. Morel, A. Micaelli, and A. Lottin. Extended-2D visual servoing. In *IEEE Int. Conf. on Robotics and Automation (ICRA'04)*, pages 267–273, New Orleans, LA, Apr. 2004.

[11] M. Costanzo, G.D. Maria, C. Natale, and Russo A. Stability and convergence analysis of 3D feature-based visual servoing. *IEEE Robotics and Automation Letters*, 7(4):12022–12029, Oct. 2022.

[12] E. Thompson. Space resection: failure cases. *The Photogrammetric Record*, 5(27):201–204, Oct. 1966.

[13] H. Michel and P. Rives. Singularities in the determination of the situation of a robot effector from the perspective view of three points. Technical Report 1850, Inria, Feb. 1993.

[14] M. Fischler and R. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, Jun. 1981.

[15] D. Oberkampf, D. Dementhon, and L.S. Davis. Iterative pose estimation using coplanar feature points. *Computer Vision and Image Understanding*, 63(3):495–511, May 1996.

[16] B. Pascual-Escudero, A. Nayak, S. Briot, O. Kermorgant, P. Martinet, M. Safey El Din, and F. Chaumette. Complete singularity analysis for the perspective-four-point problem. *Int. Journal of Computer Vision*, 129(4):1217–1237, Apr. 2021.

[17] J. Berthomieu, C. Eder, and M. Safey El Din. msolve: A library for solving polynomial systems. In *ACM Int. Symp. on Symbolic and Algebraic Computation (ISSAC'21)*, pages 51–58, Saint Petersburg, Russia, Jul. 2021.

[18] E. Malis. Improving vision-based control using efficient second-order minimization techniques. In *IEEE Int. Conf. on Robotics and Automation (ICRA'04)*, New Orleans, LA, Apr. 2004.

[19] C. Collewet, E. Marchand, and F. Chaumette. Visual servoing set free from image processing. In *IEEE Int. Conf. on Robotics and Automation (ICRA'08)*, Pasadena, CA, May 2008.

[20] S. L. Campbell and C. D. Meyer. *Generalized Inverses of Linear Transformations*. Cambridge, Mar. 2009.

[21] Hassan Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, New Jersey, 2002.

[22] C. Rucker. Integrating rotations using nonunit quaternions. *IEEE Robotics and Automation Letters*, 3(4):2979–2986, Oct. 2018.

[23] Stephen H Friedberg, Arnold J Insel, and Lawrence E Spence. *Linear algebra*. Pearson, 2014.

[24] P.H. Schönemann. A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, 31(1):1–10, Mar. 1966.

[25] W. Kabsch. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 32(5):922–923, Sep. 1976.

[26] W. Kabsch. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 34(5):827–828, Sep. 1978.

[27] G.H. Golub and C.F. Van Loan. *Matrix Computations*. J. Hopkins Uni. Press, Jan. 2013.

[28] J.C. Alexander and J.A. Yorke. The homotopy continuation method: numerically implementable topological procedures. *Trans. of the American Mathematical Society*, 242:271–284, 1978.

[29] A. Goldsztejn and L. Granvilliers. A New Framework for Sharp and Efficient Resolution of NCSP with Manifolds of Solutions. *Constraints*, 15(2):190–212, 2010.

[30] B. Martin, A. Goldsztejn, L. Granvilliers, and C. Jermann. Certified parallelotope continuation for one-manifolds. *SIAM Journal on Numerical Analysis*, 51(6):3373–3401, 2013.

[31] J. F. Canny, E. Kaltofen, and L. Yagati. Solving systems of nonlinear polynomial equations faster. In *ACM-SIGSAM Int. Symp. on Symbolic and Algebraic Computation (ISSAC'89)*, pages 121–128, 1989.

[32] B. Buchberger and F. Winkler. *Gröbner bases and applications*, volume 17. Cambridge University Press Cambridge, 1998.

[33] J. García Fontán, A. Colotti, S. Briot, A. Goldsztejn, and M. Safey El Din. Computer algebra methods for polynomial system solving at the service of image-based visual servoing. *ACM Commun. Comput. Algebra*, 56(2):36–40, nov 2022.

[34] D. Cox, J. Little, and D. Oshea. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Springer Science & Business Media, 2013.

[35] J.-C. Faugère. A new efficient algorithm for computing Gröbner bases (f4). *Journal of Pure and Applied Algebra*, 139(1-3):61–88, 1999.

[36] J.-C. Faugère. A new efficient algorithm for computing Gröbner bases without reduction to zero (f5). In *ACM Int. Symposium on Symbolic and Algebraic Computation (ISSAC'02)*, pages 75–83, 2002.

[37] J.-C. Faugère, P. Gianni, D. Lazard, and T. Mora. Efficient computation of zero-dimensional gröbner bases by change of ordering. *Journal of Symbolic Computation*, 16(4):329–344, 1993.

[38] J.-C. Faugère. Fgb: a library for computing Gröbner bases. In *Int. Congress on Mathematical Software*, pages 84–87. Springer, 2010.

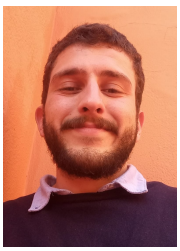
[39] Alessandro Colotti. *Stability and Regions of Attraction of Image-Based Visual Servoing Systems*. PhD thesis, Ecole Centrale de Nantes, 2023.

[40] H.-D. Chiang, M.W. Hirsch, and F.F. Wu. Stability regions of nonlinear autonomous dynamical systems. *IEEE Trans. on Automatic Control*, 33(1):16–27, 1988.

[41] B. Krauskopf, H. M. Osinga, E. J. Doedel, and O. Junge. A survey of methods for computing (un)stable manifolds of vector fields. *Int. Journal of Bifurcation and Chaos*, 15(03):763–791, 2005.

[42] Alessandro Colotti and Alexandre Goldsztejn. Practical stability and attractors of systems with bounded perturbations. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 5129–5134. IEEE, 2022.

Alessandro Colotti received the M.Sc. degree in Control Systems Engineering from the University of Padua, Padua, Italy, in 2020 and the Ph.D. degree in Control Systems Engineering from the Ecole Centrale de Nantes, Nantes, France in 2023, under the supervision of Dr. A. Goldsztejn and Dr. O. Kermorgant. He is currently an Inria postdoctoral researcher at Irisa in Rennes, France. His research interests include the stability analysis and the computation of regions of attraction for non-linear, multi-stable dynamical systems, in particular passive systems evolving on manifolds, with applications in robotics and especially image-based visual servoing.



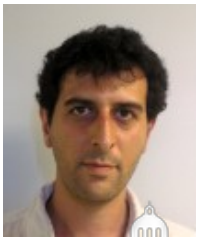
Jorge García Fontán received a Master's of Engineering degree in Aeronautical Engineering from Imperial College of London, UK, in 2019, and a Ph.D. in Computer Science from Sorbonne Université in Paris, France, in 2023. He was a Research Assistant at the LIP6 laboratory of Computer Science at Sorbonne Université for four years. His research interests include Gröbner bases and other computer algebra methods for solving polynomial systems of equations, and their applications in robotics, computer vision and, in particular, vision-based control.

Alexandre Goldsztejn received the Engineer degree in computer science and mathematics from the Institut Supérieur d'Electronique et du Numérique, Lille, France, in 2001, and the Ph.D. degree in computer science from the University of Nice Sophia Antipolis, Nice, France, in 2005. He has spent one year as a Postdoctoral Fellow with the University of Central Arkansas, Conway, AR, USA, and the University of California, Irvine, CA, USA. Since 2007, he has been a full-time CNRS Researcher with the Laboratoire des Sciences du Numeérique de Nantes, Nantes, France. His research interests include interval analysis and its applications to constraint satisfaction, nonlinear global optimization, robotics, and control.

Sébastien Briot received the B.S. and M.S. degrees in mechanical engineering in 2004 and the Ph.D. degree in robotics (under the supervision of Prof. V. Arakelian) in 2007, from the Institut National des Sciences Appliquées de Rennes, Rennes, France. He was a Postdoctoral Fellow with the Ecole de Technologie Supérieure, Montreal, QC, Canada, in 2008. In 2009, he has been recruited at CNRS as a researcher in the Laboratoire des Sciences du Numérique de Nantes, Nantes, France, where he has been the Head of the ARMEN Research Team since 2017. Since 2022, he is CNRS Director of Research in the same lab. He has authored more than 50 referred journal papers, two books, and three inventions. His research interests include the design optimization of robots and the analysis of their performance, especially their singularities. Dr. Briot received the Best Ph.D. Thesis Award in Robotics from the French CNRS in 2007. In 2011, he received two other awards: the Award for the Best Young Researcher from the French Region Bretagne and the Award for the Best Young Researcher from the French Section of the American Society of Mechanical Engineering.

François Chaumette (M'02, SM'09, F'13) is an Inria senior research scientist at Irisa in Rennes, France. He received a PhD in Computer Science from the University of Rennes 1 in 1990. His research interests include robotics and computer vision, especially visual servoing and active perception. He supervised more than 30 PhD students and published over 300 journal or conference papers, with the 2002 Best IEEE Transactions on Robotics and Automation Paper Award, the 2019 Best IEEE Robotics and Automation Letters Paper Award, and the 2020 Best IEEE Robotics and Automation Magazine Paper Award. He was Founding Senior Editor of the IEEE Robotics and Automation Letters, member of the Editorial Board of the Int. Journal of Robotics Research, and Senior Editor of the IEEE Transactions on Robotics.

Olivier Kermorgant graduated from École Centrale Paris, France in 2004. From 2008 to 2011 he was with the Lagadic group at Inria Rennes where he received the Ph.D. degree in signal processing from University of Rennes, France in 2011. He then joined the Ocean Systems Laboratory at Heriot-Watt University, Edinburgh, Scotland as a Research Assistant. He was Assistant Professor at University of Strasbourg, France from 2012 to 2014. Since 2015 he has been Associate Professor at Centrale Nantes and LS2N, France. His research interests include sensor-based robot control, maritime robotics and optimization.



Mohab Safey El Din defended his PhD in 2001 at Univ. Pierre et Marie Curie. He is Professor at Sorbonne Université since 2012 and a honorary member of Institut Universitaire de France. His research interests lie in computer algebra and polynomial system solving and their application in sciences, especially in robotics. He is an editorial board member of the Journal of Symbolic Computation, and of the Journal of Algebra (computational section). He served several times in the program committee of the International Symposium on Symbolic and Algebraic Computation – the premier conference in this area – which he chaired in 2017. With his collaborators, he received several distinguished paper awards.