On the Structure of the Inverse Kinematics Map of a Fragment of Protein Backbone

R.J. Milgram, Guanfeng Liu[†], and J.C. Latombe[†]
Department of Mathematics, Stanford University
milgram@math.stanford.edu
[†] Department of Computer Science, Stanford University
{liugf,latombe}@cs.stanford.edu

Abstract

Loop closure in proteins requires computing the values of the inverse kinematics (IK) map for a backbone fragment with $2n \geq 6$ torsional degrees of freedom (dofs). It occurs in a variety of contexts, *e.g.*, structure determination from electron-density maps, loop insertion in homology-based structure prediction, backbone tweaking for protein energy minimization, and study of protein mobility in folded states. The first part of this paper analyzes the global structure of the IK map for a fragment of protein backbone with 6 torsional dofs and a slightly idealized kinematic model, called the canonical model. This model, which assumes that every two consecutive torsional bonds C_α–C and N–C_α are exactly parallel, makes it possible to separately compute the inverse orientation map and the inverse position map. The singularities of both maps and their images, the critical sets, respectively decompose $SO(3)$ and $\mathbb{R}^3$ into open regions where the number of IK solutions is constant. This decomposition leads to a constructive proof of the existence of a region in $\mathbb{R}^3 \times SO(3)$ where the IK of the 6-dof fragment has 16 solutions. The second part of this paper extends this analysis to study fragments with more than 6 torsional dofs. It describes an efficient recursive algorithm to sample IK solutions for such fragments, by identifying the feasible range of each successive torsional dof. A numerical homotopy algorithm is then used to deform the IK solutions for a canonical fragment into solutions for a non-canonical fragment. Computational results for fragments ranging from 8 to 30 dofs are presented.

Keywords: Protein backbone, inverse kinematics, loop closure, singularity analysis, conformation space sampling, homotopy algorithm.

1 Introduction

A protein is a long sequence of amino-acids connected by peptide bonds [6]. Under normal physiological conditions, bond lengths and angles, as well as dihedral angles around peptide bonds are usually assumed constant at their equilibrium values. A protein is then modeled as a serial articulated linkage – the protein's backbone – with short side-chains protruding from it. In this paper, we will focus on the serial linkage formed by the backbone. Each amino-acid contributes three atoms – N, C_α, and C – and two torsional degrees of freedom (*dof*s) to it. These dofs correspond to changes in the values of the dihedral angles around the N–C_α and the C_α–C bonds. See Figure 1. These angles are commonly referred to as the ϕ and ψ angles,

respectively. Hence, the backbone of a protein formed by r amino-acids has $2r$ dofs. The study of the kinematics of a protein backbone – *i.e.*, the study of its intrinsic mobility without regard to the forces needed to actually create motion – has attracted considerable interest (*e.g.*, [5, 12, 13, 14, 25]), in particular to compute protein conformations subject to geometric constraints. Of special interest is the inverse kinematics (IK) problem for a backbone fragment. This problem is also known as the *loop closure problem* [2, 5, 12, 24].

Let F be a backbone fragment with $2n$ dihedral angles $\theta = (\theta_1, \cdots, \theta_{2n})$, $n \geq 3$, where each θ_i is either a ϕ or a ψ angle. One terminus (usually a pair of atoms) of F is considered fixed and the other mobile. We attach a Cartesian frame W (for "world") to the fixed terminus and another frame T (for "tool") to the mobile terminus. The *forward kinematics* of F is the function f that maps each value θ to the *pose* (position and orientation) of T relative to W. The *inverse kinematics* f^{-1} of F maps every pose c of T to the corresponding values of θ. The IK problem is defined as follows: given c, compute the values of θ. In this paper, we consider the two cases where $n = 3$ and $n > 3$.

Case where $n = 3$: It is well known that for a 6-dof fragment F the number of values of $f^{-1}(c)$ is finite and upper bounded by 16 (as they are the zeros of a monovariate polynomial of degree 16) [5, 13]. However, it has recently been questioned whether this bound is tight [5, 8, 24]. Algorithms are available to compute the values of f^{-1} when $n = 3$, for instance the elegant analytical method developed in [5], but they only return the values of f^{-1} for *given* poses of T.

Each pose of T is a point in the 6-D manifold $\mathbb{R}^3 \times SO(3)$, where $SO(3)$ is the group of 3×3 rotation matrices. One goal of this paper is to analyze the *global* structure of f^{-1} over $\mathbb{R}^3 \times SO(3)$. This leads us to study the singularities of the forward kinematics function f, that is, the values of θ where the Jacobian of f looses rank. The set of singularities of f (the *singular set* of f) maps to a set of *critical* poses of T (the *critical set* of f). According to the Morse-Sard theorem [11, 16], the critical set of f decomposes $\mathbb{R}^3 \times SO(3)$ into *regular* regions, such that over each such region f^{-1} keeps a constant number of values.

However, any decomposition of a 6-D space into regions can be extremely complex. So, here we proceed as follows, with a slightly idealized kinematic model, which we call the *canonical* model of F. First, we break the forward kinematics map into the position and the orientation maps. Next, we show that the orientation map has a relatively simple inverse, which leads to a simple region decomposition of the 3-D manifold $SO(3)$. Finally, we establish the inverse position map on the restricted domain of θ defined by the inverse orientation map. For each possible orientation of T, we decompose $\mathbb{R}^3$ into regular regions over which the structure of the inverse position map remains constant. As we will see, the critical set of the position map turns out to be an unexpectedly complicated 2-D surface in $\mathbb{R}^3$.

A key outcome of this study is that, when $n = 3$, f^{-1} actually has 16 values over at least one small region of $\mathbb{R}^3 \times SO(3)$. Moreover, the way this result is obtained can be used to construct poses of T in this region. This outcome is not completely surprising, since the conditions under which the IK of a 6-dof serial linkage has less than 16 solutions have been established in [15, 19] and a 6-dof fragment of a protein backbone does not satisfy any of them.

Case where $n > 3$: When F contains more than 6 dihedral angles, it can deform while keeping its two termini fixed. For a given reachable pose c of T, the conformational space of

F is then a $(2n - 6)$-D manifold, which is often called the *self-motion* manifold of F at c [1].

Several applications in biology, *e.g.*, loop insertion in protein structure prediction with homology modeling [10], backbone tweaking for energy minimization [21], analysis of loop and backbone mobility in folded states [4, 20], and structure determination from blurred electron-density maps [23], require sampling conformations from self-motion manifolds. Some existing sampling methods, like the Cyclic Coordinate Descent [2] and the Jacobian's pseudo-inverse [23], first pick the torsional angles in F at random and then "close" the fragment by iteratively reducing the distance between the current pose of T and its goal pose c. But these iterative methods tend to be rather slow and do not guarantee a good coverage of the self-motion manifold. A more direct method, RLG (Random Loop Generator) [4], successively samples $2n - 6$ torsional angles θ_1 through θ_{2n-6} in F and then uses an IK method to compute the remaining 6 angles. To pick the value of each angle θ_i, $i = 1, ..., 2n - 6$, RLG first computes a conservative approximation of the range of values of θ_i that enable T to reach the pose c, given the values already selected for $\theta_1, ..., \theta_{i-1}$; next, it samples the value of θ_i from the approximated range. However, RLG's approximation is very conservative. It not only ignores orientation reachability, but position reachability is also loosely approximated by bounding the volumes reachable by the successive atoms in F with spherical shells. In practice, RLG has a high failure rate, *i.e.*, many choices of torsional angles eventually prevent T from reaching the goal pose.

The other goal of this paper is to develop an efficient algorithm to sample self-motion manifolds. This algorithm is based on a generalization of our analysis of the IK map f^{-1} when $n = 3$ to longer fragments F. Like RLG, our method successively picks $2n - 6$ torsional angles. However, it considers both positional and orientational reachability, and computes a tight approximation of the feasible range of each angle, given the values already sampled for the previous angles. Although the mathematical derivation of this approximation is not straightforward, its computation is very fast. Computational tests show that our new algorithm samples self-motion manifolds with a reasonably good success rate.

This paper is organized as follows. In Section 2 we present the canonical kinematic model of a $2n$-dof protein fragment, along with the notation used in the rest of the paper. In particular, we define the forward and inverse kinematics maps, and we decompose them into position and orientation maps. Next, in Sections 3–5, we consider the case where $n = 3$. In Sections 3 and 4, we analyze the structures of the inverse orientation and position maps, respectively, in that case. In Section 5, we establish the existence of a region in $\mathbb{R}^3 \times SO(3)$ where f^{-1} has 16 distinct values. We show one instance of a pose of T in this region, along with its 16 inverse kinematics solutions. In Section 6, we extend our IK analysis to study longer backbone fragments ($n > 3$) with non-canonical kinematics. This extension leads to an efficient algorithm for sampling the self-motion manifold of a $2n$-dof fragment F when $n > 3$. A numerical homotopy algorithm is used to deform the IK solutions sampled for a canonical fragment into solutions for a non-canonical fragment.
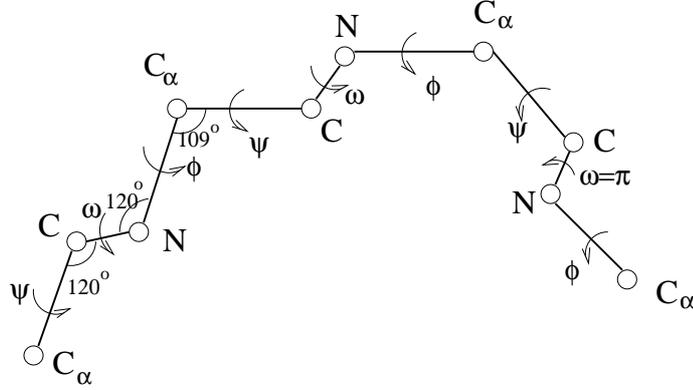
Figure 1: Linkage model of a 6-dof fragment of a protein backbone

2   Kinematic Model of a Protein Fragment

2.1   Linkage model

Consider a $2n$-dof fragment F of a protein backbone (Figure 1), in which all non-constant dihedral angles (ϕ and ψ) are comprised between two C_α atoms. We denote the successive atoms of F by C_α^1, C^1, N^2, C_α^2, ..., C^n, N^{n+1}, C_α^{n+1}. We assume that bond lengths and bond angles are constant, and that the dihedral angle ω around each peptide bond is constant and equal to π [6]. The $2n$ dofs of F correspond to the dihedral angles ϕ around the bonds N^i–C_α^i, for $i = 2, \ldots, n + 1$ and the dihedral angles ψ around the bonds C_α^i–C^i, for $i = 1, \ldots, n$. In the following, we denote the angle ϕ around N^i–C_α^i by θ_{2i-2} and the angle ψ around C_α^i–C^i by θ_{2i-1}. So, we parameterize each conformation of F by $\theta = (\theta_1, \ldots, \theta_{2n})$. So, each angle θ_i with an even (resp. odd) subscript is a ϕ (resp. ψ) angle.

We idealize the kinematic model of F slightly more, by assuming that every two consecutive torsional bonds C_α^i–C^i and N^{i+1}–C_α^{i+1} are exactly parallel. This corresponds to assuming equal bond angles $\angle(C_\alpha^i C^i N^{i+1})$ and $\angle(C^i N^{i+1} C_\alpha^{i+1})$. We also assume that bond lengths and bond angles have the same values in all residues. We refer to this idealized model as the *canonical* model of F.

For convenience, but with no additional simplification, we then replace the representation of Figure 1 by the kinematically equivalent model of Figure 2, in which F is a sequence of n identical units, each made of two perpendicular links, a "long" one of length ℓ_2 and a "short" one of length ℓ_1. We number the links $1, 2, \ldots, 2n$, so that each link $2i - 1$ is a long link and each link $2i$ is a short link. Each long link $2i - 1$ originates at atom C_α^i, and each short link $2i$ ends at atom C_α^{i+1}. Angle θ_{2i-1} rotates short link $2i$ about long link $2i - 1$. So, each short link moves in a plane perpendicular to the preceding long link. Angle θ_{2i} rotates long link $2i + 1$ about an axis parallel to long link $2i - 1$ and passing through the extremity of short link $2i$. Link $2i + 1$ makes a constant angle $\alpha = \angle(N^{i+1} C_\alpha^{i+1} C^{i+1}) - 90°$ with the plane perpendicular to link $2i - 1$.

We add a long link $2n + 1$ at the end of F. This is the only link in F whose orientation

4
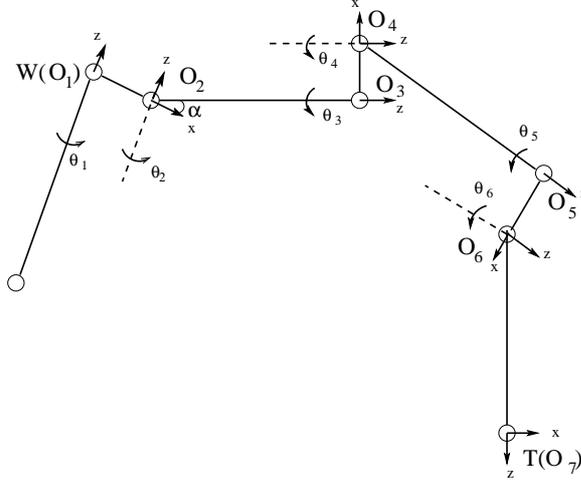
Figure 2: Canonical model of a 6-dof backbone fragment

depends on θ_{2n}. However, the torsional angle around it is not considered to be part of F.

The specific values chosen for the bond angles $\angle(C_\alpha^i C^i N^{i+1})$ and $\angle(C^i N^{i+1} C_\alpha^{i+1})$ and for α are not important for the rest of this paper. They can easily be changed in our software. The computational results shown in this paper (except those presented in Section 6.2, where we extend our methods to non-canonical fragments) have been obtained with $\angle(C_\alpha^i C^i N^{i+1})$ and $\angle(C^i N^{i+1} C_\alpha^{i+1})$ set to $120°$, α set to $19°$ (hence, $\angle(N^i C_\alpha^i C^i) = 109°$ for all i), and ℓ_1 and ℓ_2 respectively set[1] to 1.1518Å and 3.665Å.

2.2 Forward kinematics

We attach a Cartesian frame to each link, as shown in Figure 2. The origin of frame $2i - 1$ is positioned at the extremity of long link $2i - 1$, with its z-axis pointing along this link. The origin of frame $2i$ coincides with the extremity of short link $2i$, its z-axis is parallel to the z-axis of frame $2i - 1$ and its x-axis points along the short link $2i$. Each angle θ_i rotates the frame $i + 1$ with respect to the frame i, hence has no effect on the orientation of the frame i. So, we consider frame 1 as fixed and we select it as the "world" frame W. We select frame $2n + 1$ as the "tool" frame T.

We represent the pose of frame i relative to W by a position vector, $p_i \in \mathbb{R}^3$, the origin of this frame in W, and a rotation matrix O_i, its orientation in W. The rotation matrix is an element of the group $SO(3)$ defined by:

$$SO(3) := \{A \in \mathbb{R}^{3\times 3} \mid A^T A = I_{3\times 3},\ \det(A) = 1\}.$$

---

[1] The lengths ℓ_1 and ℓ_2 are derived from the standard bond lengths 1.47Å, 1.53Å, and 1.33Å for the bonds N-C_α, C_α-C, and C-N, respectively.

We have:

$$\begin{aligned}
O_1 &= I_{3\times 3}, \\
O_{2i} &= O_{2i-1}R_{2i-1}, \\
O_{2i+1} &= O_{2i}R_{2i}L,
\end{aligned}$$

where:

$$R_i = \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) & 0 \\ \sin(\theta_i) & \cos(\theta_i) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad L = \begin{bmatrix} -\sin(\alpha) & 0 & \cos(\alpha) \\ 0 & -1 & 0 \\ \cos(\alpha) & 0 & \sin(\alpha) \end{bmatrix}.$$

Notice that $L^2 = I_{3\times 3}$, hence $L^{-1} = L$. Moreover $L^T = L$.

Thus

$$\begin{aligned}
O_{2i+1} &= R_{1;2}LR_{3;4}L\cdots LR_{2i-1;2i}L, \\
O_{2i} &= R_{1;2}L\cdots LR_{2i-1},
\end{aligned}$$

where $R_{i;j} = R_iR_j$.

The *forward kinematics* of F is the function f that maps each $\theta \in (S^1)^{2n}$, where S^1 denotes the unit circle, to the pose $f(\theta)$ of T relative to W, where $f(\theta) \in \mathbb{R}^3 \times SO(3)$. This map is the product (p, ρ) of a position map p:

$$p : (S^1)^{2n} \to \mathbb{R}^3, \qquad \theta \to \sum_{i\leq n} O_{2i+1}v_1 + \sum_{i\leq n} O_{2i}v_2,$$

where $v_1 = [0, 0, \ell_2]^T$ and $v_2 = [\ell_1, 0, 0]^T$, and an orientation map ρ:

$$\rho : (S^1)^{2n} \to SO(3), \qquad \theta \to O_{2n+1}.$$

2.3 Inverse kinematics

This paper studies the structure of the *inverse kinematics* map $f^{-1} = (p, \rho)^{-1}$ that maps *points* of $\mathbb{R}^3 \times SO(3)$ to *subsets* of $(S^1)^6$. Noticing that for any $(X, R) \in \mathbb{R}^3 \times SO(3)$:

$$(p, \rho)^{-1}(X, R) = p^{-1}(X) \cap \rho^{-1}(R),$$

we proceed in two steps. First, we derive the inverse orientation map $\rho^{-1} : SO(3) \to (S^1)^6$ mapping points of $SO(3)$ to subsets of $(S^1)^6$ of the form $\Sigma \times (S^1)^3$, with Σ a finite set. Next, we compute the inverse position map $p_M^{-1}(X)$, where p_M is the forward position map p with its domain restricted to the manifold $M = \sigma \times (S^1)^3$ for individual points $\sigma \in \Sigma$. The next three sections only consider the case where $n = 3$.

3 Inverse Orientation Map ($n = 3$)

In this section, we study the structure of the inverse orientation map ρ^{-1} over $SO(3)$ when $n = 3$. We first reduce the number of parameters of ρ to obtain a reduced orientation map $\hat{\rho}$. We then study the singular set of $\hat{\rho}$ and its image, the critical set. The main outcome of this section is that, for every non-critical orientation R of T, $\rho^{-1}(R)$ is either empty or is the disjoint union of two 3-D tori. For every critical orientation R, $\rho^{-1}(R)$ forms a single 3-D torus or, in one special case, a 4-D torus.

3.1   Reduction

It follows from Subsection 2.2 that the forward orientation map ρ of a 6-dof fragment is:

$$\rho : (S^1)^6 \to SO(3), \qquad \theta \to R_{1;2} L R_{3;4} L R_{5;6} L.$$

We notice that the 6 dihedral angles $\theta_1, ..., \theta_6$ appear by pairs $\theta_{2i-1} + \theta_{2i}$. So, we write $\tau_i = \theta_{2i-1} + \theta_{2i}$, $i = 1,\ 2,\ 3$, and $\tau = (\tau_1, \tau_2, \tau_3)$. When θ runs over $(S^1)^6$, τ runs over the 3-D torus $(S^1)^3$.

The orientation map ρ factors as the composition of two maps denoted $(+)$ and $\hat{\rho}$:

$$\rho = \hat{\rho} \circ (+) : (S^1)^6 \to (S^1)^3 \to SO(3)$$

where:

$$(+) : (S^1)^6 \to (S^1)^3, \qquad (\theta_1, ..., \theta_6) \to (\theta_1 + \theta_2, \theta_3 + \theta_4, \theta_5 + \theta_6)$$

and

$$\hat{\rho} : (S^1)^3 \to SO(3), \qquad \tau \to R_{\tau_1} L R_{\tau_2} L R_{\tau_3} L,$$

with:

$$R_{\tau_i} \;\; = \;\; \begin{bmatrix} \cos(\tau_i) & -\sin(\tau_i) & 0 \\ \sin(\tau_i) & \cos(\tau_i) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Given $R \in SO(3)$, the values of $\hat{\rho}^{-1}(R)$ are the solutions of:

$$\hat{\rho}(\tau) := R_{\tau_1} L R_{\tau_2} L R_{\tau_3} L \;\; = \;\; R.$$

Since L is a constant matrix with $L = L^{-1}$, this equation is equivalent to:

$$\hat{\rho}(\tau) L := R_{\tau_1} L R_{\tau_2} L R_{\tau_3} \;\; = \;\; RL. \tag{1}$$

We notice that the matrix $\hat{\rho}(\tau) L$ defines the orientation O_6 of frame 6. Since the z-axis of frame 6 is independent of τ_3, we further reduce Equation (1) by eliminating the variable τ_3, To do this, we define the action:

$$A_z : SO(3) \to S^2, \qquad R \to Rz,$$

where $z = [0,\ 0,\ 1]^T$. A_z maps any rotation matrix R to its last column. Since every column of R is a unit vector, $A_z(R) \in S^2$, where S^2 denotes the unit 2-D sphere. We remark that $A_z(R_{\tau_3}) = z$. So, applying A_z to both sides of Equation (1) yields the reduced equation:

$$A_z(\hat{\rho}(\tau) L) := R_{\tau_1} L R_{\tau_2} L z \;\; = \;\; RLz. \tag{2}$$

We can solve this equation for (τ_1, τ_2). The value of τ_3 is then uniquely determined by the equation:

$$R_{\tau_3} = (R_{\tau_1} L R_{\tau_2} L)^T RL. \tag{3}$$

To each solution $\tau = (\tau_1, \tau_2, \tau_3)$ of Equations (2) and (3) corresponds a set of values of $\theta = (\theta_1, ..., \theta_6)$ such that $\theta_{2i-1} + \theta_{2i} = \tau_i$ for $i = 1,\ 2,\ 3$. This set is a 3-D torus $(S^1)^3$.

3.2 Singular set

The singularities of $\hat{\rho}$ are the points in $(S^1)^3$ where the 3×3 Jacobian matrix $J\hat{\rho}$ has rank less than 3. The Jacobian of $\hat{\rho}$ is defined as $J\hat{\rho} = [D_{\tau_1}\hat{\rho}, D_{\tau_2}\hat{\rho}, D_{\tau_3}\hat{\rho}]\hat{\rho}^{-1}$, where $D_{\tau_i}\hat{\rho}$ is the matrix $\left[\frac{\partial \hat{\rho}_{jk}}{\partial \tau_i}\right]$. $J\hat{\rho}$ maps the tangent space $\mathcal{T}(S^1)^3$ to the Lie algebra $so(3)$ [17]. In Appendix A, with respect to a given basis of $so(3)$, we compute:

$$J\hat{\rho} \;=\; [z, \; R_{\tau_1}Lz, \; R_{\tau_1}LR_{\tau_2}Lz].$$

We easily verify that $J\hat{\rho}$ has at least rank 2 and that it has exactly rank 2 if and only if:

$$det(J\hat{\rho}) \;=\; \sin(\tau_2)\cos(\alpha) \;=\; 0.$$

Since $\cos(\alpha) \neq 0$, the (rank-2) *singular set* of $\hat{\rho}$ is the union $\{\tau \mid \tau_2 = 0\} \cup \{\tau \mid \tau_2 = \pi\}$.

This result has a simple intuitive geometric interpretation. The orientation of frame T relative to W is the composition of three rotations of angles τ_1, τ_2, and τ_3 around directions parallel to long links 1, 3, and 5, respectively. These links are coplanar if and only if $\tau_2 = 0$ or π. When this is the case, no infinitesimal changes of τ_1, τ_2, and τ_3 can rotate T around a direction perpendicular to the plane containing links 1, 3, and 5. So, the resulting infinitesimal changes of the orientation of T only span a 2-D subspace of the 3-D tangent space $\mathcal{T}SO(3)$. Instead, whenever links 1, 3, and 5 are non-coplanar, infinitesimal changes of τ_1, τ_2, and τ_3 result in changes of the orientation of T that span all dimensions of $\mathcal{T}SO(3)$.

3.3 Critical set and number of solutions

Consider the map:

$$\eta : (S^1)^3 \to S^2, \qquad \tau \to A_z(\hat{\rho}(\tau)L) := R_{\tau_1}LR_{\tau_2}Lz,$$

which appears in the left-hand side of Equation (2). It has the same singular set $\{\tau \mid \tau_2 = 0\} \cup \{\tau \mid \tau_2 = \pi\}$ as $\hat{\rho}$. Therefore, its *critical set* – *i.e.*, the image by η of this singular set – is the union of two sets C_1 and C_2 (see Figure 3):

- C_1 is the subset of S^2 spanned by $R_{\tau_1}LR_{\tau_2}Lz$ when $\tau_2 = 0$ and τ_1 varies over S^1. Since $(R_{\tau_1}LR_{\tau_2}Lz)_{\tau_2=0} = R_{\tau_1}LLz = R_{\tau_1}z = z$, C_1 consists of a single point, z, which corresponds to the situation where the z-axes of W and frame 6 are parallel. Indeed, when $\tau_2 = 0$, the z-axis of frame 6 is parallel to the z-axis of W for any value of τ_1 (recall that the dihedral angle around every peptide bond is equal to π).

- C_2 is the subset of S^2 spanned by $R_{\tau_1}LR_{\tau_2}Lz$ when $\tau_2 = \pi$ and τ_1 varies over S^1. We have: $(R_{\tau_1}LR_{\tau_2}Lz)_{\tau_2=\pi} = [(\sin(2\alpha)\cos(\tau_1), \; \sin(2\alpha)\sin(\tau_1), \; -\cos(2\alpha)]^T$. So, C_2 is the circle perpendicular to the z-axis and passing through the point $LR_\pi Lz$.

The Morse-Sard theorem, together with the compactness of $(S^1)^3$, tells us that the inverse map η^{-1}, hence $\hat{\rho}^{-1}$, is a fibration with a compact differentiable manifold as fiber in each of the two open subsets of S^2 bounded by C_1 and C_2. Likewise the transitive S^1 action on C_2
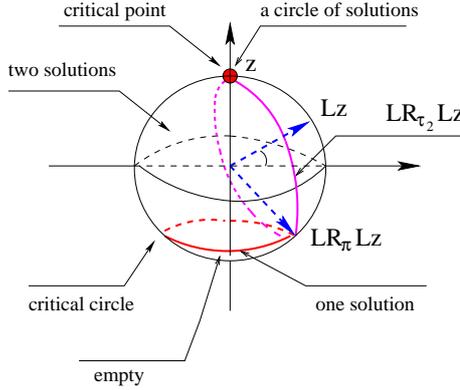
Figure 3: Critical set of η

given by R_{τ_1} shows that η restricted to $\eta^{-1}(C_2)$ is also a fibering over C_2. To determine the exact structure of η^{-1} in these four regions, we notice that:

$$L(LR_{\tau_2}Lz) = [\cos(\alpha)\cos(\tau_2),\ \cos(\alpha)\sin(\tau_2),\ \sin(\alpha)]^T.$$

So, $LR_{\tau_2}Lz$ is a circle perpendicular to Lz that is fully contained in the subset of S^2 between C_1 and C_2, except at $\tau_2 = 0$ and $\tau_2 = \pi$ where it coincides with C_1 and C_2, respectively (see Figure 3). For any fixed $\tau_1 \in S^1$, the set $R_{\tau_1}LR_{\tau_2}Lz$ is the circle obtained by rotating $LR_{\tau_2}Lz$ by τ_1 around the z axis. Therefore, for every point s in the region between C_1 and C_2, the circle $R_{\tau_1}LR_{\tau_2}Lz$ contains s for two distinct values of τ_1. Hence, η^{-1} has two values (τ_1^k, τ_2^k), $k = 1,\ 2$. In C_1, $s = z$ and the set of values of $\eta^{-1}(s)$ is $\{(\tau_1, 0)\ |\ \tau_1 \in S^1\}$. For any $s \in C_2$, $\eta^{-1}(s)$ has a single value of the form (τ_1, π). When s is "below" C_2, the set of values of $\eta^{-1}(s)$ is empty.

Corresponding to each value (τ_1, τ_2) of $\eta^{-1}(s)$ there is a unique value of τ_3 given by Equation (3), hence a single value of $\hat{\rho}^{-1}(R)$.

Thus, for every $R \in SO(3)$ such that the orientation of the z-axis of O_6 is in the subset of S^2 between C_1 and C_2, the set of values of $\rho^{-1}(R)$ is the disjoint union of two 3-D tori. For every R such that the orientation of the z-axis of O_6 is in C_2, $\rho^{-1}(R)$ consists of a single 3-D torus. For every R such that the orientation of the z-axis of O_6 is in C_1, $\rho^{-1}(R)$ is a 4-D torus. For all other R, $\rho^{-1}(R)$ is empty. On each 3-D torus of $\rho^{-1}(R)$, every $\theta_{2j-1} + \theta_{2j}$, $j = 1,\ 2,\ 3$, is constant.

4   Inverse Position Map

We now study the structure of the inverse position map p^{-1} for a given orientation $R \in SO(3)$ of the frame T such that $\rho^{-1}(R)$ is neither empty nor a 4-D torus, *i.e.*, consists of one or two 3-D tori. Let M be one of these tori. We consider the position map p_M restricted to M. Rather than studying the singular set of p_M, we directly determine the critical set of p_M by studying the positions $X \in \mathbb{R}^3$ where the number of values of p_M^{-1} changes. We then compute the decomposition of $\mathbb{R}^3$ by this critical set.

9

4.1 Restriction to $\rho^{-1}(R)$

From Subsection 2.2 we derive the forward position map p for a 6-dof protein fragment:

$$p : (S^1)^6 \to \mathbb{R}^3, \qquad \theta \to (R_{1;2}L + R_{1;2}LR_{3;4}L + R_{1;2}LR_{3;4}LR_{5;6}L)v_1$$
$$+ (R_1 + R_{1;2}LR_3 + R_{1;2}LR_{3;4}LR_5)v_2.$$

Since all $\tau_j = \theta_{2j-1} + \theta_{2j}$, $j = 1, 2, 3$, are constant on M, each point on M is uniquely defined by the values of θ_1, θ_3, and θ_5. Hence:

$$p_M : (S^1)^3 \to \mathbb{R}^3, \qquad (\theta_1, \theta_3, \theta_5) \to v_M + (R_1 + R_{\tau_1}LR_3 + R_{\tau_1}LR_{\tau_2}LR_5)v_2,$$

where $v_M = (R_{\tau_1}L + R_{\tau_1}LR_{\tau_2}L + R_{\tau_1}LR_{\tau_2}LR_{\tau_3}L)v_1$ is a constant vector. We notice that $R_1 v_2$, $R_{\tau_1}LR_3 v_2$, and $R_{\tau_1}LR_{\tau_2}LR_5 v_2$ are three circles that have the same radius ℓ_1, but lie in three different planes (except when $\tau_2 = 0$). In any case, the image of $(S^1)^3$ by p_M is the Minkowski sum of these three circles translated by v_M.

Given any $X \in \mathbb{R}^3$, solving the inverse kinematics $p_M^{-1}(X)$ amounts to finding the solutions of the equation:

$$X' = LR_{-2}v_2 + R_3 v_2 + R_{\tau_2}LR_5 v_2. \tag{4}$$

where $X' = LR_{\tau_1}^T(X - v_M)$ and R_{-2} represents the rotation of $-\theta_2$ around z. To further simplify the right-hand side of (4), we multiple both sides of the equation by $R_{\pi/2}$, the rotation of $\pi/2$ around z, which leads to introducing the map $\hat{p}_M$ as shown below:

$$X'' = \hat{p}_M(-\theta_2, \theta_3, \theta_5) := R_{\pi/2}(LR_{-2} + R_3 + R_{\tau_2}LR_5)v_2 \tag{5}$$

where $X'' = R_{\pi/2}X'$. Notice that in (5) the second circle is contained in the xy-plane, while the other two circles are arranged symmetrically with respect to the xz-plane when $\tau_2 = \pi$.

4.2 Critical set of $\hat{p}_M$

We now directly determine the critical positions X'' by computing where the number of values of $\hat{p}_M^{-1}$ changes.

For convenience, we rewrite Equation (5) by replacing $-\theta_2$, θ_3, and θ_5 by $-\pi/2+u$, $-\pi/2+t$, and $-\pi/2+w$, respectively, and the constant τ_2 by $\gamma+\pi$. We also normalize ℓ_1 to 1 (ℓ_2 becomes 3.1820 correspondingly) without affecting the structure of the fragment's conformation space, so that all three circles in the equation now have normalized radius 1. We get:

$$X'' - r(t) = q(u, w) \tag{6}$$

with

$$r(t) = \begin{bmatrix} c_t \\ s_t \\ 0 \end{bmatrix}, \qquad q(u, w) = \begin{bmatrix} c_u + c_\gamma c_w - s_\alpha s_\gamma s_w \\ -s_\alpha s_u + s_\alpha c_\gamma s_w + s_\gamma c_w \\ c_\alpha (s_u + s_w) \end{bmatrix}, \tag{7}$$

where c_* and s_* stand for $\cos(*)$ and $\sin(*)$.
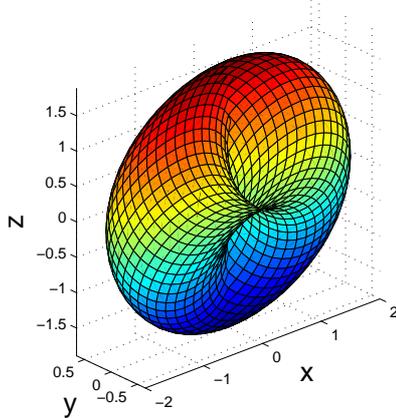
Figure 4: The quartic surface Q

For any given $X'' \in \mathbb{R}^3$, when t varies, the left-hand side of Equation (6), $X'' - r(t)$, spans a unit circle centered at X''. When u and w vary, the right-hand side, $q(u, w)$, spans a bounded and connected quartic surface Q (see Figure 4) that is the Minkowski sum of the first and the last circles in Equation (5). Appendix B gives a detailed description of Q. So, Equation (6) can be solved by computing the intersection points between the circle $X'' - r(t)$ and the cross-section curve of the surface Q by the plane containing this circle.

The unit vector normal to this plane is z. So, the equation of the plane containing $X'' - r(t)$ can be written as:

$$z^T q = d, \tag{8}$$

where $d = z^T X''$ is the last component of X''. We let P_d denote the plane defined by this equation. When X'' varies over $\mathbb{R}^3$, P_d translates, but its orientation remains fixed.

By replacing q by the expression $q(u, w)$ in (7) in Equation (8), we get the equation of the cross-section Q_d of Q by P_d in terms of u and w:

$$s_u + s_w = \frac{d}{c_\alpha}. \tag{9}$$

For any X'', the solutions to Equation (6) are obtained by computing the intersection points between between Q_d and the circle $X'' - r(t)$. This amounts to solving for the (real) roots of an 8th-order polynomial (see Appendix B). The number of real solutions is always even when the multiplicity of the roots is taken into account, and varies between 0 and 8 as X'' varies over $\mathbb{R}^3$.

Let $d_{\min}$ and $d_{\max}$ be the two extreme values of d, between which the plane P_d intersects the surface Q. For any $d \in [d_{min}, d_{max}]$, the values of X'' such that the circle $X'' - r(w)$ lies in P_d and is tangent to Q_d form a curve $\mathcal{X}_d$, called the *discriminant curve* at d. This discriminant curve is the locus of all points where the 8th-order polynomial has at least one multiple real root. The 2-D surface formed by the union of all the discriminant curves as d runs over the interval $[d_{\min}, d_{\max}]$ is the critical set $\mathcal{X}$, i.e., $\mathcal{X} = \bigcup_{d \in [d_{\min}, d_{\max}]} \mathcal{X}_d$. Figures 5 and 6

11
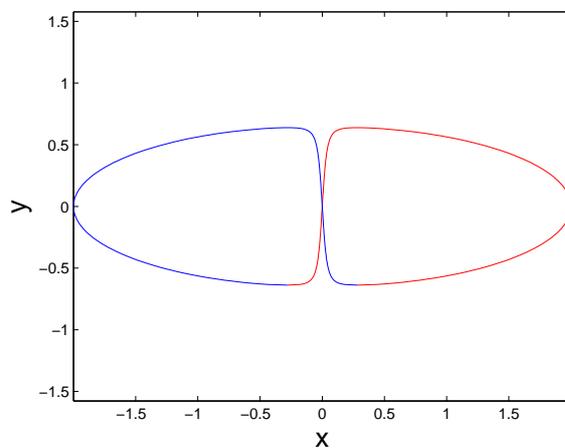
Figure 5: The cross-section of Q in the plane P_d for $\gamma = 0$ and $d = 0.0378$
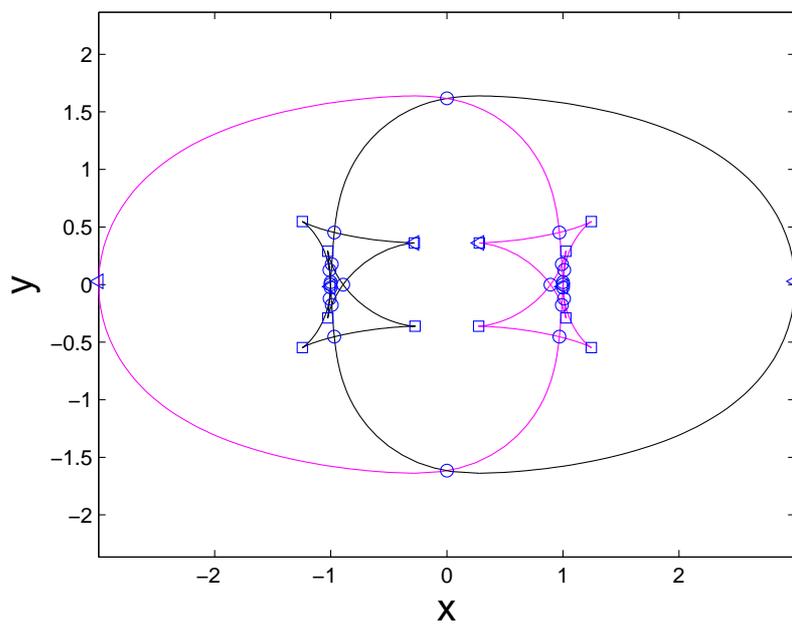
Figure 6: The discriminant curve $\mathcal{X}_d$ for $\gamma = 0$ and $d = 0.0378$. The centers of the small squares are the cusp points of $\mathcal{X}_d$, the centers of the small circles are its self-intersection points, and the centers of the small triangles are the vertical tangency points. A zoom on a portion of $\mathcal{X}_d$ is shown in Figure 8-(a)

12
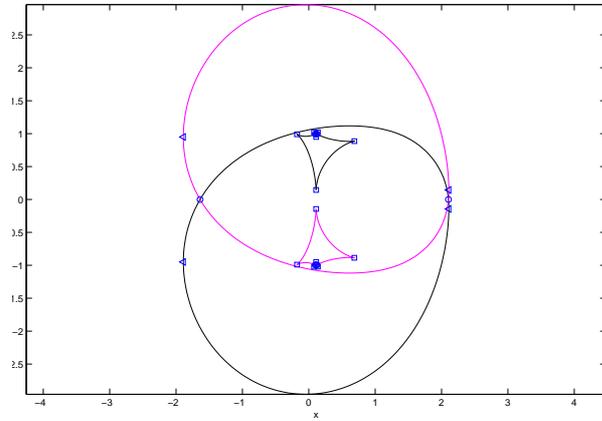
Figure 7: The discriminant curve in another plane. A zoom on a portion of this curve is shown in Figure 8-(b).
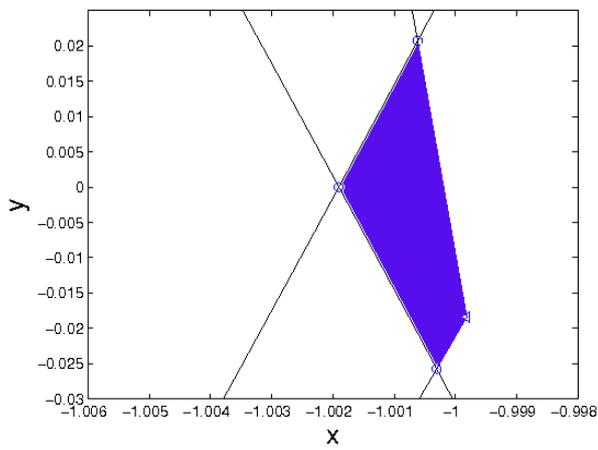
(a)
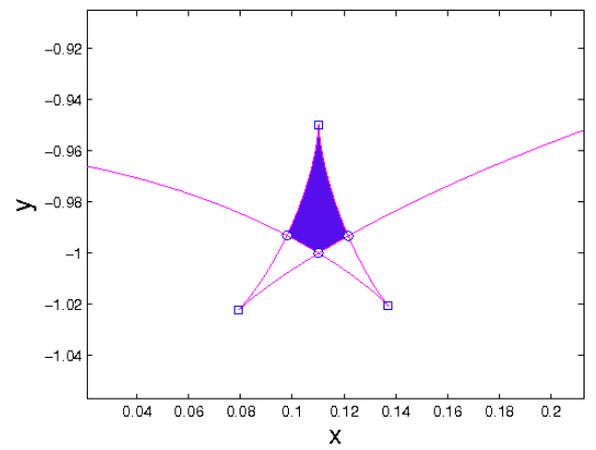
(b)

Figure 8: Enlarged portions of Figure 6-(a) and Figure 7-(b) with an 8-solution region shown in blue in each case.
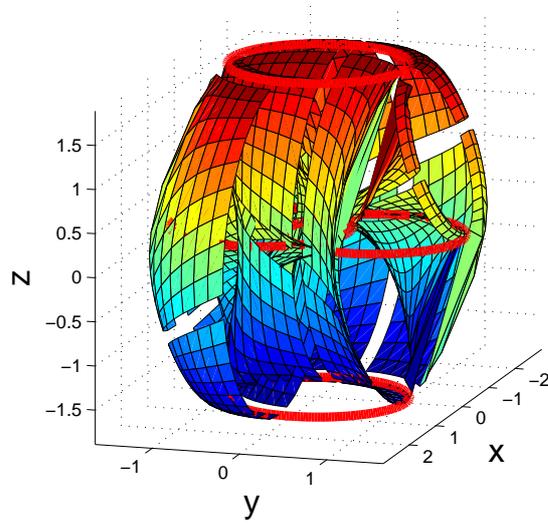
Figure 9: Graphic rendering of $\mathcal{X}$ for $\gamma = 0$ (one patch of $\mathcal{X}$ is not rendered in order to show the interior of $\mathcal{X}$)
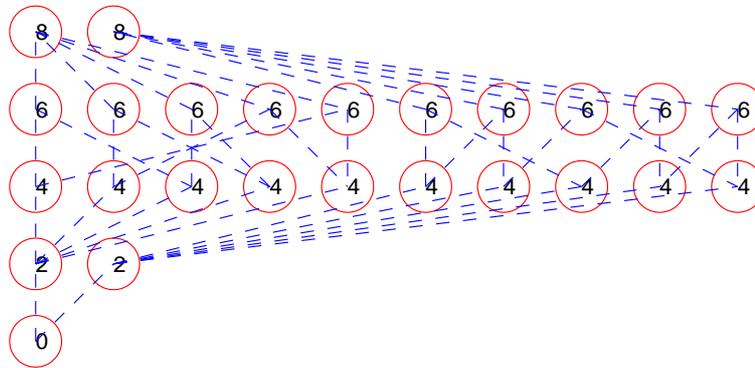
Figure 10: The planar region graph determined by the discriminant curve of Figure 6. The circles depict regions and the dashed lines connect adjacent regions. The number of solutions of the inverse position map in each region is shown in the corresponding circle.

14

show a cross-section of Q. and the corresponding discriminant curve, respectively. Figure 7 displays another discriminant curve. Figure 8 contains enlarged portions of Figures 6 and 7. These figures show that a discriminant curve can be surprisingly complicated, with several cusp points and self-intersection points. The graphic rendering of $\mathcal{X}$ in Figure 9 further reveals the complexity of $\mathcal{X}$. (An animation of the cross-section of Q and the corresponding discriminant curve when d varies is available at `www.stanford.edu/~phwu1/curve` in the case where $\gamma = 0$.)

4.3   Decomposition of $\mathbb{R}^3$ into regions

The critical set $\mathcal{X}$ decomposes $\mathbb{R}^3$ into open 3-D regions such that the inverse position map has a constant number of values over each region. To compute this decomposition, we first compute the decomposition of a plane P_d by $\mathcal{X}_d$. Next, we partition $[d_{min}, d_{max}]$ into open intervals, such that over each such interval the discriminant curves $\mathcal{X}_d$ are homotopic to each other, *i.e.*, keep the same combinatorial structure determined by the cusp points and self-intersection points. The decomposition of $\mathbb{R}^3$ is obtained by "stacking" the planar decompositions computed in the successive intervals.

4.3.1   Decomposition of P_d

We use a classical plane-sweep algorithm [18] that sweeps a line L parallel to the y-axis across the plane P_d from left to right. This algorithm constructs a set S of sub-regions and their adjacency relation. At the start of the sweep, S is initialized to the empty set. During the sweep, whenever L crosses a cusp point, a self-intersection point, or a vertical tangency point, sub-regions are added to S and the adjacency relation is updated. When the sweep is completed, adjacent sub-regions in S that are not separated by $\mathcal{X}_d$ are merged together into regions that form the decomposition of P_d. The outcome is a planar *region graph* in which the nodes are the computed regions and the edges represent the adjacency relation (*i.e.*, any two nodes connected by an edge represent two regions that are adjacent along a portion of $\mathcal{X}_d$). The number of solutions of the inverse position map increases or decreases by 2 whenever one crosses the boundary between two adjacent regions.

Computing the cusp and self-intersection points analytically is difficult. So, we approximate the discriminant curve into a polygonal curve made of short segments and we compute those points numerically using this approximation. Whenever the polygonal approximation is sufficiently accurate (*e.g.*, the maximum distance between the line segments and the discriminant curve is less than $0.5\epsilon > 0$, where ϵ is the minimum of the radii of all regions), the computed graph correctly predicts all regions and their adjacency relationship. The number of distinct solutions in each region is obtained by picking any point X'' in this region (not too close from its boundary) and solving Equation (6) at this point.

Figure 10 shows the region graph computed from the discriminant curve displayed in Figures 6 and 8-(a). The nodes are drawn as circles and the edges as dashed lines. The number of IK solutions in each region is given in the corresponding circle. An animation of the discriminant curve and the corresponding region graph when d varies is available at `www.stanford.edu/~phwu1/curve` in the case where $\gamma = 0$.
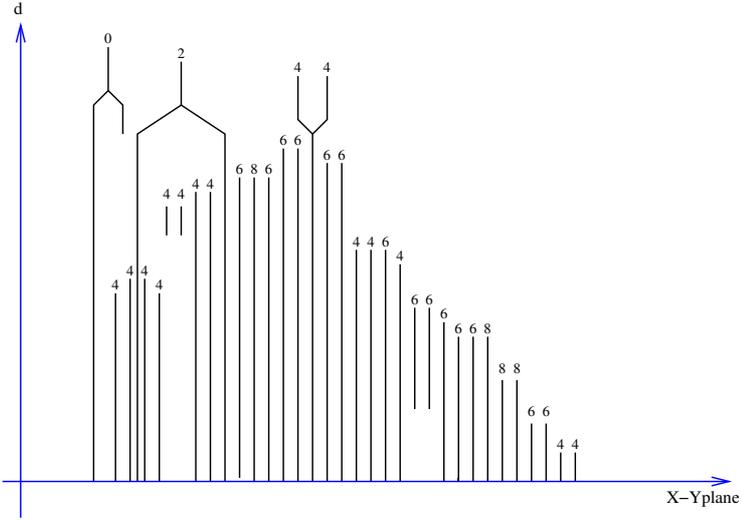
Figure 11: The region graph for the upper half-space of $\mathbb{R}^3$ ($d \in [0, \infty)$) where $\gamma = 1.3$. The region graph for the lower half-space is center-symmetric to the upper half-space one. Each connected tree or line segment denotes a region. The number of IK solutions for each region is shown above each line segment.

4.3.2 Decomposition of $\mathbb{R}^3$

When d varies from $-\infty$ to $+\infty$, the planar region graph in P_d changes only at a finite number of *critical* values of d, which we denote d_i, $i = 1, ..., m$. Over each open interval (d_i, d_{i+1}), $i = 0, ..., m$, with $d_0 = -\infty$ and $d_{m+1} = +\infty$, the discriminant curves $\mathcal{X}_d$ are homotopic and so the region graph remains constant.

Let G_i be the planar region graph in interval (d_i, d_{i+1}). The region decomposition of $\mathbb{R}^3$ is obtained by merging every pair of regions from G_i and G_{i+1}, for all $i = 0, ..., m$, that are adjacent, but not separated by $\mathcal{X}$. The corresponding nodes of the planar region graphs are also merged to obtain the region graph of the decomposition of $\mathbb{R}^3$. The region graph for the case where $\gamma = 1.3$ is shown in Figure 11. In this example $\mathbb{R}^3$ is partitioned into 39 regions including the one where there is no IK solution. The remaining issue is to determine the m critical values $d_1, ..., d_m$ of d. The critical set $\mathcal{X}$ (see Figure 9) is a 2-D surface made of smooth patches separated by cusp and self-intersection curves. The cusp (resp. self-intersection) curves are the locus $\mathcal{X}^{\text{cusp}}$ (resp. $\mathcal{X}^{\text{self}}$) of all the cusp (self-intersection) points of the discriminant curves $\mathcal{X}_d$ when d varies from $-\infty$ to $+\infty$. The critical values of d are contributed either by $\mathcal{X} \backslash (\mathcal{X}^{\text{cusp}} \cup \mathcal{X}^{\text{self}})$, or by $\mathcal{X}^{\text{cusp}}$, or by $\mathcal{X}^{\text{self}}$.

The critical values contributed by $\mathcal{X} \backslash (\mathcal{X}^{\text{cusp}} \cup \mathcal{X}^{\text{self}})$ are simply the values of d such that the plane P_d is tangent to $\mathcal{X}$. It is not difficult to prove that this case occurs only when P_d is tangent to the surface Q. Equation (9) of Q_d yields the 3 critical values $-2c_\alpha$, 0, $2c_\alpha$, which define four critical circles on $\mathcal{X}$ (shown in red in Figure 9):

$$\{(t, u, w) \in (S^1)^3 \mid c_u = c_w = 0, \, t \in S^1\}.$$

The computation of the critical values contributed by $\mathcal{X}^{\text{cusp}}$ and $\mathcal{X}^{\text{self}}$ is much more complicated. Currently we perform this computation numerically.

16

5   Existence of a 16-Solution Region

Theorem 1 *There exists a nonempty open region in $\mathbb{R}^3 \times SO(3)$ such that for all (X, R) in this region, $(p, \rho)^{-1}(X, R)$ contains 16 points.*

Proof: Consider first an orientation $R_0 \in SO(3)$ that lies in the critical circle C_2 shown in Figure 3. The set $\rho^{-1}(R_0)$ is a copy of $(S^1)^3$. There is a non-empty open region $E_0 \subset C_2$ such that for all R in E_0, $p(\rho^{-1}(R))$ contains an open region U so that, for any $X \in U$, $p^{-1}(X)$ contains 8 points, as shown in Figure 8. Let R' be a non-critical orientation that is close to E_0. Then $\rho^{-1}(R')$ is a disjoint union of two 3-D tori M_k, $k = 1, 2$. For each p_{M_k}, there exists a nonempty open region E_{M_k} with 8 inverse image points. Moreover, for R' sufficiently close to E_0, $E = E_1 \cap E_2$ is nonempty. Then $(p, \rho)^{-1}(X, R')$ has 16 solutions for all $X \in E$. ∎

Using the idea in the proof, we constructed the following pose (X, R) of T:

$$X = \begin{bmatrix} 2.0563 \\ 4.9057 \\ -2.3925 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} 0.6742 & -0.3715 & -0.6383 \\ 0.2378 & -0.7091 & 0.6638 \\ -0.6992 & -0.5993 & -0.3897 \end{bmatrix},$$

such that $(p, \rho)^{-1}(X, R)$ contains 16 solutions (for a fragment in which $\ell_1 = 1$ and $\ell_2 = 3.1820$). These are shown in Figures 12. It is easily seen that the existence of 16-solution region is independent of the link lengths as long as the short links all have the same length.

6   Generalization

In this section, we first extend the results of the previous sections to propose a recursive algorithm that randomly samples the conformation space of a canonical protein fragment F with $2n > 6$ torsional dofs, given a fixed pose of the frame T relative to W. This algorithm partitions F into a $(2n - 6)$-dof fragment and the remaining 6-dof fragment. Values of the angles τ_i and θ_{2i-1} in the $(2n-6)$-dof fragment are sampled from their feasible ranges computed using the orientation and position equations. The remaining 6 torsional angles are determined using the results of Sections 3 and 4. In the second part of the section, we present a numerical homotopy algorithm that deforms conformations sampled for a canonical fragment with $2n > 6$ dofs into conformations of a non-canonical fragment with the same number n of residues and the same two anchors. So, the combination of these two algorithms allows us to sample the conformation space of a non-canonical fragment, given a fixed pose of the frame T relative to W.

6.1   Inverse kinematics of a long canonical fragment

Following the same general approach as in Sections 3 and 4, we successively consider the inverse orientation map and the inverse position map.

We first propose a recursive method that successively samples the angles $\tau_i = \theta_{2i-1} + \theta_{2i}$, $i = 1, 2, ..., n-3$. For each angle τ_i, we first compute its feasible range, *i.e.*, the interval of values for which the inverse orientation map is non-empty, given the values previously sampled for $\tau_1, \cdots, \tau_{i-1}$. The value of τ_i is sampled from this interval. The last three angles $(\tau_{n-2}, \tau_{n-1}, \tau_n)$ are computed using the results established in Section 3.
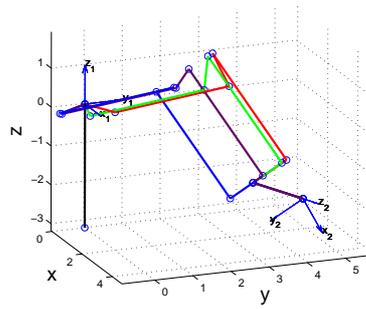
(a)
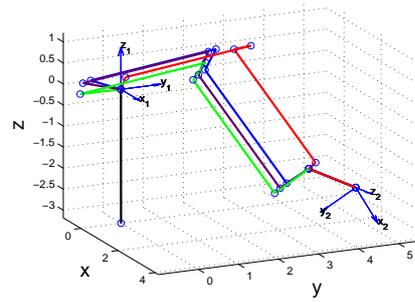
(b)
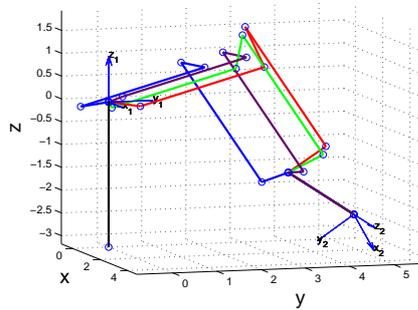
(c)
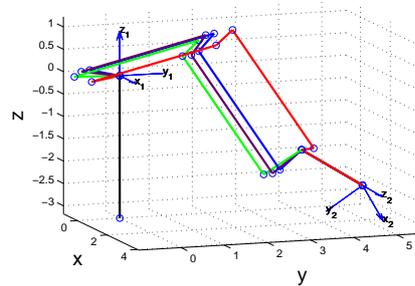
(d)

Figure 12: Sixteen IK solutions (see Section 5): each of the four drawings shows four conformations.

Then we propose a similar method to sample the values of the angles $\theta_1, \cdots, \theta_{2n-7}$. Again, the last three angles $(\theta_{2n-5}, \theta_{2n-3}, \theta_{2n-1})$ are computed using the results of Section 4. Each pair of values for τ_i and θ_{2i-1} determines the value of θ_{2i}.

Sampling the values of τ_i: Like in Section 3, we define the reduced orientation map $\hat{\rho}$ by:

$$\hat{\rho} : (S^1)^n \to SO(3), \qquad (\tau_1, \cdots, \tau_n) \to R_{\tau_1} L R_{\tau_2} L \cdots L R_{\tau_n} L.$$

Given $R \in SO(3)$, the values of $\hat{\rho}^{-1}(R)$ are the solutions $(\tau_1, ..., \tau_n)$ of the equation:

$$R_{\tau_1} L R_{\tau_2} L \cdots L R_{\tau_n} L = R.$$

After eliminating τ_n, we get the following equation:

$$R_{\tau_1} L R_{\tau_2} L \cdots L R_{\tau_{n-1}} L z = R L z. \tag{10}$$

The map $\eta \ : \ (\tau_1, ..., \tau_n) \in (S^1)^n \to R_{\tau_1} L R_{\tau_2} L \cdots L R_{\tau_{n-1}} L z \in S^2$ is singular if and only if each of the angles τ_i, for $i = 2, ..., n-1$, is equal to 0 or π, *i.e.*, if all torsional axes are coplanar. So, the critical set on S^2 is the union of 2^{n-2} critical circles defined by:

$$\{R_{\tau_1} L R_{\tau_2} L \cdots L R_{\tau_{n-1}} L z \mid \tau_2 = 0 \text{ or } \pi \ ; \cdots \ ; \tau_{n-1} = 0 \text{ or } \pi\}.$$

Because $LL = I$, some of these circles coincide, so that the number of distinct critical circles is only $\lceil n/2 \rceil$ (the smallest integer greater than or equal to $n/2$). Furthermore, when n is odd, one circle reduces to the point z. All critical circles have the same normal vector z, so that they partition S^2 into parallel regions.

It is not difficult to verify that when $n > 3$, the solution set of Equation (10) is non-empty in all critical circles and all parallel regions between these circles. So, the image of η covers the whole sphere S^2.

Therefore, the values of all the angles τ_i, for $i = 1, ..., n-3$, can be sampled uniformly at random from the interval $[-\pi, \pi)$. The following remaining equation (whose unknowns are τ_{n-3}, τ_{n-2}, and τ_{n-1}) has a non-empty solution set:

$$R_{\tau_{n-3}} L R_{\tau_{n-2}} L R_{\tau_{n-1}} L z = L R^T_{\tau_{n-4}} L \cdots L R^T_{\tau_1} R L z, \tag{11}$$

since the image of its left-hand side covers the whole S^2. Then, given the values sampled for $\tau_1, ..., \tau_{n-4}$, the feasible range of τ_{n-3} is the set of all values of τ_{n-3} such that the following equation (whose unknowns are τ_{n-2}, and τ_{n-1}) has a non-empty solution set:

$$R_{\tau_{n-2}} L R_{\tau_{n-1}} L z = L R^T_{\tau_{n-3}} L \cdots L R^T_{\tau_1} R L z. \tag{12}$$

Following Section 3, the critical set of the left-hand side of this equation consists of a critical point and a critical circle that decompose S^2 into two parallel open regions, such that the bottom one has zero IK solutions. The right-hand side, $L R_{\tau_{n-3}} L \cdots L R^T_{\tau_1} R L z$, spans a circle in S^2 when τ_3 varies over S^1. So, we compute the intersection points of this latter circle with the critical circle (see Appendix C). This gives two intervals for τ_{n-3}, $(\tau^1_{n-3}, \tau^2_{n-3})$ and $(\tau^2_{n-3}, \tau^1_{n-3} + 2\pi)$, of which one is the feasible range of values of τ_{n-3}.

It then remains three angles $(\tau_{n-2}, \tau_{n-1}, \tau_n)$, corresponding to a 6-dof fragment, for which the structure of the inverse orientation map has been completely determined in Section 3.3.

It follows directly from the above presentation that the feasible region for $(\tau_1, \tau_2, \cdots, \tau_{n-3})$ is connected. Thus, the set of all conformations of F that achieve a given orientation R of the frame T is either connected or consists of two disconnected components. The topological structure of $\eta^{-1}(RLz)$ can be derived following the method presented in [22] for planar closed chains with revolute joints.

Sampling the values of θ_{2i-1}: Following Section 4, for a fixed value of $(\tau_1, \cdots, \tau_n)$, we can write the position map as follows:

$$p(\theta_1, \theta_3, \cdots, \theta_{2n-1}) = v_0 + (R_1 + R_{\tau_1}LR_3 + \cdots + R_{\tau_1}L \cdots LR_{\tau_{n-1}}LR_{2n-1})v_2, \tag{13}$$

where v_0 is a constant vector. So, $p(\theta_1, \theta_3, \cdots, \theta_{2n-1})$ spans the Minkowski sum of n circles that have the same radius ℓ_1, but lie in n different planes in general. Given $X \in \mathbb{R}^3$, the values of $p^{-1}(X)$ are the solutions $(\theta_1, ..., \theta_{2n-1})$ of the equation:

$$(R_1 + R_{\tau_1}LR_3 \cdots + R_{\tau_1}L \cdots LR_{\tau_{n-1}}LR_{2n-1})v_2 = X - v_0. \tag{14}$$

We move θ_1 in the right-hand side of this equation to get:

$$(R_3 + R_{\tau_2}LR_5 \cdots + R_{\tau_2}L \cdots LR_{\tau_{n-1}}LR_{2n-1})v_2 = LR_{-\tau_1}(X - v_0) - LR_{-\tau_1}R_1v_2, \tag{15}$$

which has the same general form as Equation (14). The inverse image $p^{-1}(X)$ is the set of all the solutions of Equation (15) for all $\theta_1 \in S^1$.

We note that, for any given value of θ_1, the right-hand side of Equation (15) defines a point Y in $\mathbb{R}^3$. The value of θ_1 is feasible if and only if Y is contained in the volume V spanned by the Minkowski sum of the $n-1$ circles on the left-hand side of the equation. When θ_1 varies over S^1, Y spans a circle C. The intersection of C with V determines the feasible range of θ_1. However, computing this intersection exactly is difficult. Instead, we notice that the shape of the volume spanned by the Minkowski sum of three (or more) non-coplanar circles is approximately convex. So, we bound V by its convex hull $\mathcal{H}(V)$ and we compute the intersection of C with $\mathcal{H}(V)$. We conservatively approximate the feasible range of θ_1 to be the interval of values that define the portion of C contained in $\mathcal{H}(V)$.

The intersection of C with $\mathcal{H}(V)$ is calculated using the second-order cone programming method [3]. Appendix D proves that the convex hull $\mathcal{H}(V)$ is the Minkowski sum of $n-1$ unit disks defined by the $n-1$ circles. Let $\tilde{x}_i$ and $\tilde{y}_i$ be, respectively, the two unit base vectors for the planes of the ith circle. To form the convex constraints, C is approximated by a regular polygon with 16 (or more) sides:

$$C = \{X_j\lambda_1 + X_{j+1}\lambda_2 \mid \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1, j = 1, \cdots, 16\},$$

where X_j, $j = 1, \cdots, 16$, are the vertices of the regular polygon and $X_{17} = X_1$. We solve[2] the following 16 second-order cone programming problems for the intersections between the

---

[2]In our implementation, we use Sedumi, version 1.1, which is available from http://sedumi.mcmaster.ca.

(a)   (b)   (c)
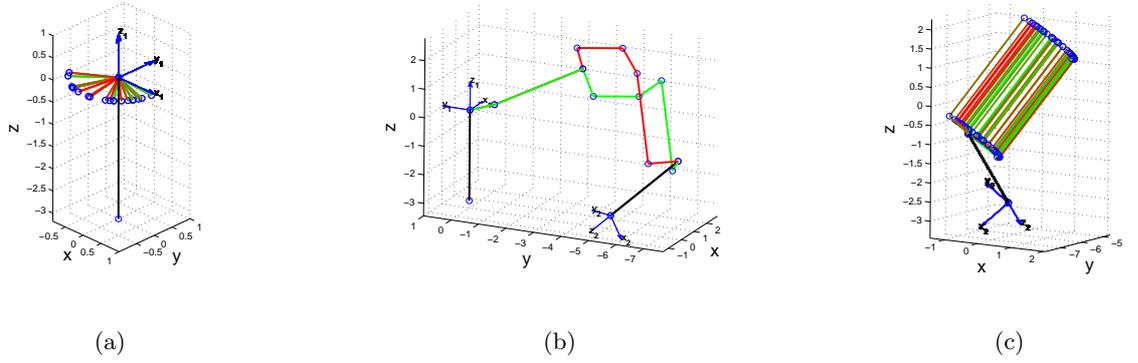
Figure 13: Sampled conformations of an 8-dof fragment: (a) The proximal end; (b) Two conformations; (c) The distal end.

(a)   (b)   (c)
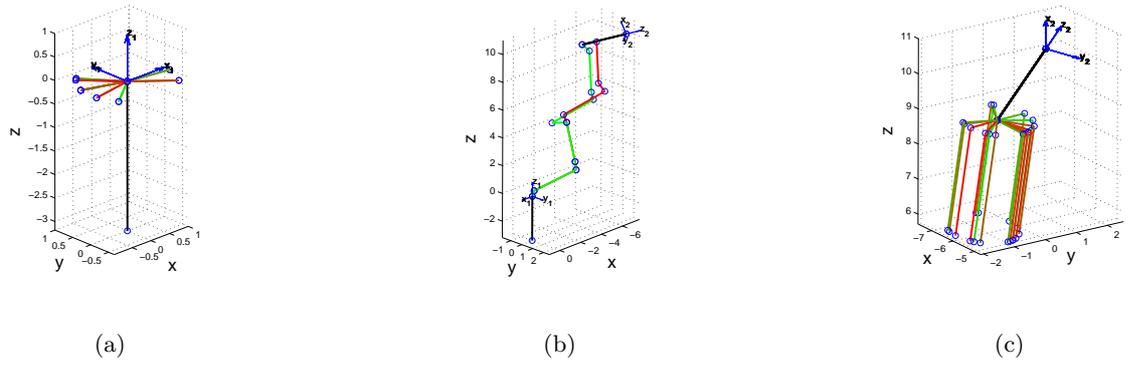
Figure 14: Sampled conformations of a 10-dof fragment: (a) The proximal end; (b) Two conformations; (c) The distal end.

(a)   (b)   (c)
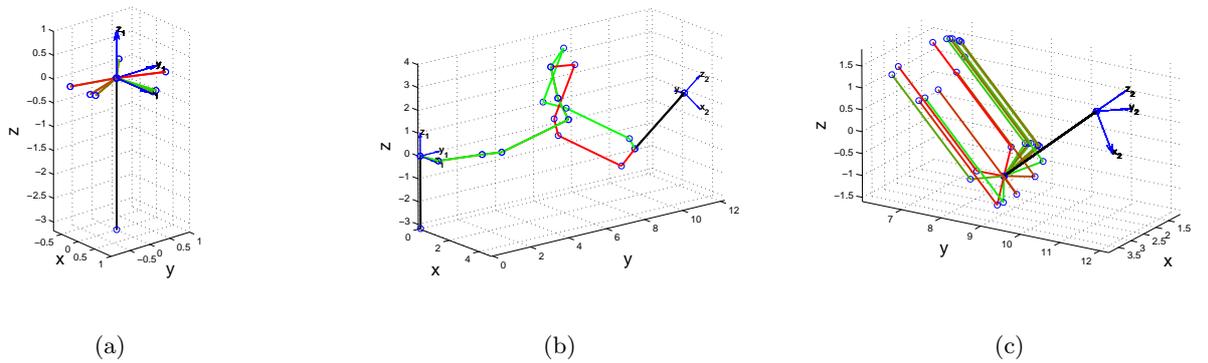
Figure 15: Sampled conformations of a 12-dof fragment: (a) The proximal end; (b) Two conformations; (c) The distal end.

21

regular polygon and the boundary of $\mathcal{H}(V)$:

$$\max \lambda_2$$

$$\text{subject to}: \quad X_j\lambda_1 + X_{j+1}\lambda_2 = \begin{bmatrix} \tilde{x}_1, \tilde{y}_1, \cdots, \tilde{x}_{(n-1)}, \tilde{y}_{(n-1)} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ \vdots \\ x_{(n-1)} \\ y_{(n-1)} \end{bmatrix}$$

$$x_i^2 + y_i^2 \le z_i^2, \, i = 1, \cdots, n-1$$
$$1 - z_i \ge 0$$
$$z_i \ge 0$$
$$\lambda_i \ge 0, \, i = 1,2$$
$$\lambda_1 + \lambda_2 = 1.$$

The feasible range of θ_1 is identified by checking if a point within the open interval between two consecutive candidate intersections is feasible.

For any given value of θ_1 sampled in its (approximated) feasible range, we can similarly compute the (approximated) feasible range of θ_3, then pick a value of θ_3 in this range, and so on. The recursion ends when the left-hand side of Equation (15) is the sum of three circles. The value of $(\theta_1, \theta_3, \cdots, \theta_{2n-7})$ determines at most 8 solutions for $(\theta_{2n-5}, \theta_{2n-3}, \theta_{2n-1})$ (as discussed in Section 4.2 and Appendix B).

Computational results Figures 13, 14, and 15 show conformations sampled with the recursive algorithms described above for fragments with 8, 10, and 12 dofs, respectively. Each angle is sampled uniformly at random in its computed feasible range. These sampled conformations cover well the self-motion space of each of the three fragments.

Our algorithm is not guaranteed to succeed in sampling a conformation of the fragment at each trial. There are two distinct reasons for this. First, the feasible ranges for the angles τ_i are computed without taking the inverse position map into account. Therefore, a value of τ_i picked in its computed feasible range may not allow the frame T to reach its goal position. Second, the feasible ranges computed for the angles θ_{2i-1} are slightly conservative due to the approximation of a Minkowski sum of circles by its convex hull.

We measured the failure rate of our IK algorithm for values of n (number of residues) ranging from 4 to 15. For each value of n, we made $N = 300$ independent trials. Each trial consisted of picking a new pose of the frame T at random and running our IK algorithm on this pose. Each pose of T was obtained by sampling the angles θ_i, $i = 1, ..., 2n$ uniformly at random from $[0, 2\pi)$ and then computing the forward kinematics map. For each pose of T, the IK algorithm returned failure if a computed feasibility range was empty or there was no solution for the last three angles $(\tau_{n-2}, \tau_{n-1}, \tau_n)$ or $(\theta_{2n-5}, \theta_{2n-3}, \theta_{2n-1})$. We estimated the failure rate of the IK algorithm as the ratio of the number of failures by N, and the average computation time of a sampled conformation as the ratio of the total running time of the N trials by the number of successful trials. The results are plotted in Figure 16 and 17, respectively.[3] Although the total failure rate may look rather high (between 80 and 90%) at

---

[3]Our software is written in Matlab and runs on a P3(1.13-GHz) under WindowsXP.
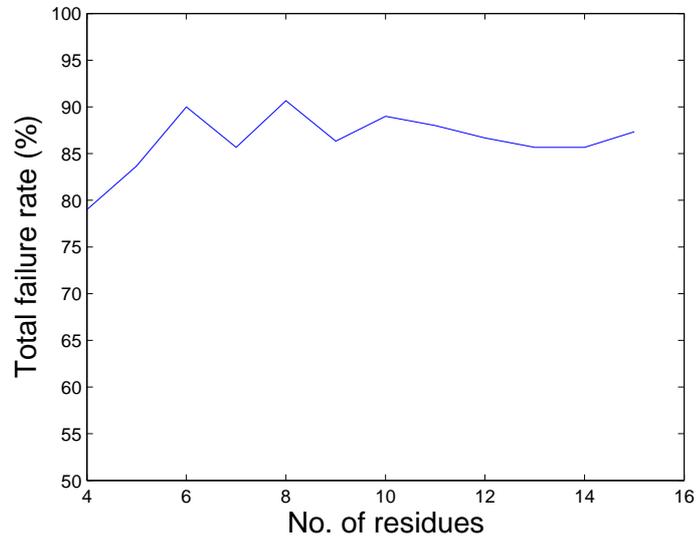
Figure 16: Total failure rate
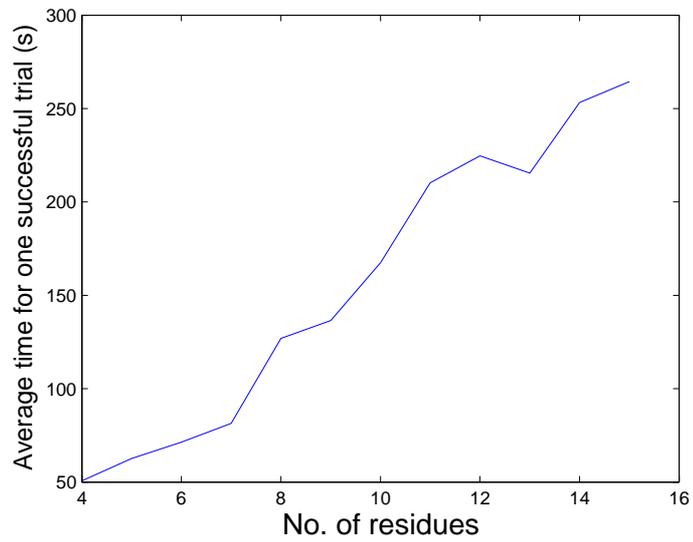
Figure 17: Average computation time for a successful trial

first sight, it is much smaller than with RLG and still yields a good running time per sampled conformation. In fact, a success rate of 10 to 20% to sample a submanifold of dimensionality $2n - 6$ by sampling each one of the $2n$ parameters defining the ambient manifold is a fairly good result. Note also that the success rate varies little as n increases.

Additional tests (not shown here) indicates that the approximation of the Minkowski sums of multiple circles by their convex hulls is responsible for relatively few failures. Most failures are caused by the choice of the angles τ_i that does not take the inverse position map into account.

We also implemented and tested a combination of our algorithm and the RLG algorithm. The combined algorithm computes the feasible ranges of the angles τ_i by intersecting the intervals computed by our algorithm with those obtained by applying the RLG algorithm to Equation (13), in which v_0 is treated not as a constant, but as a function of τ_i. On fragments of various lengths, the combined algorithm did not achieve a better success rate than our algorithm alone. In most cases, the intervals returned by RLG are supersets of the intervals computed by our algorithm.

6.2 Inverse kinematics of a long non-canonical fragment

All results presented so far assume a canonical fragment in which every two successive bonds $C_\alpha^i - C^i$ and $N^{i+1} - C_\alpha^{i+1}$ are exactly parallel. Moreover, all computational tests have been conducted with models where the angle between each two successive bonds $N^i - C_\alpha^i$ and $C_\alpha^i - C^i$ is 109 degrees. We now describe a method to deform a conformation for a canonical fragment into a conformation for a non-canonical fragment with the same number of dofs.

When successive bonds $C_\alpha^i - C^i$ and $N^{i+1} - C_\alpha^{i+1}$ are not exactly parallel and the angle between each two successive bonds $N^i - C_\alpha^i$ and $C_\alpha^i - C^i$ is not exactly 109 degrees, the model shown in Figure 2 must be adapted. More specifically, the rotation axis corresponding to θ_{2i} (the z-axis of O_{2i}) is then obtained by rotating that of θ_{2i-1} (the z-axis of O_{2i-1}) by an angle $\zeta_i \neq 0$ along the y-axis of O_{2i-1}. The lengths of the links must also be modified slightly. The non-canonical model returns to canonical when $\zeta_i = 0$ and $\alpha_i = 19$ degree.

Although the principle of the method described below is general, we simplify our presentation by assuming that the deviations relative to the canonical model are the same for all $i = 1, ..., n$. With this simplification, we let $(\zeta^{\text{goal}}, \alpha^{\text{goal}}) \neq (0, 19\pi/180)$ denote the parameters for the non-canonical fragment. Similarly, $(\zeta^{\text{init}}, \alpha^{\text{init}}) = (0, 19\pi/180)$ stands for the parameters for the canonical model.

In the product space $\mathbb{R}^2 \times \mathbb{R}^{2n}$ of the parameters and joint angles, the IK problem is to find solutions for:

$$R_1 R^0 R_2 L R_3 R^0 R_4 L \cdots L R_{2n-1} R^0 R_{2n} L = R \qquad (16)$$

$$R_1 v_2 + R_1 R^0 R_2 L v_1 + \cdots + R_1 R^0 \cdots R_{2n-1} v_2 + R_1 R^0 \cdots R^0 R_{2n} L v_1 = X \qquad (17)$$

for a given pose $(X, R) \in \mathbb{R}^3 \times SO(3)$, where the matrix $R^0 = \text{Rot}(y, \zeta)$ represents the rotation of angle ζ about the y-axis. Let θ^{init} be a solution to this problem when $(\zeta, \alpha) = (\zeta^{\text{init}}, \alpha^{\text{init}})$. Such a solution can be sampled using the methods described in Subsection 6.1. We then apply the numerical homotopy algorithm [9] described below to derive a solution θ^{goal} for $(\zeta, \alpha) = (\zeta^{\text{goal}}, \alpha^{\text{goal}})$.

Differentiating Equations (16) and (17) yields:

$$J_1 d\theta + \mathcal{V}_1 \begin{bmatrix} d\zeta \\ d\alpha \end{bmatrix} = 0$$

and

$$J_2 d\theta + \mathcal{V}_2 \begin{bmatrix} d\zeta \\ d\alpha \end{bmatrix} = 0,$$

where J_1 and $\mathcal{V}_1$ (resp. J_2 and $\mathcal{V}_2$) are the Jacobian matrices for the orientation map (resp. position map). Thus, we get:

$$d\theta = - \begin{bmatrix} J_1 \\ J_2 \end{bmatrix}^\dagger \begin{bmatrix} \mathcal{V}_1 \\ \mathcal{V}_2 \end{bmatrix} \begin{bmatrix} d\zeta \\ d\alpha \end{bmatrix}, \tag{18}$$

where $[\cdot]^\dagger$ denotes the pseudo inverse of $[\cdot]$. We consider the straight path $t \in [0,1] \to (\zeta(t), \alpha(t)) \in \mathbb{R}^2$ such that $(\zeta(0), \alpha(0)) = (\zeta^{\text{init}}, \alpha^{\text{init}})$ and $(\zeta(1), \alpha(1)) = (\zeta^{\text{goal}}, \alpha^{\text{goal}})$ and we integrate $d\theta$ along this path to obtain the IK solution for the non-canonical fragment:

$$\theta^{\text{goal}} = \theta^{\text{init}} + \int_0^1 d\theta.$$

Here we use a two-step integration algorithm. The first step tracks the solution of Equations (16)-(17) along the path by computing:

$$
\begin{aligned}
{}^0\theta^{i+1} &= \theta^i + d\theta, \\
\zeta^{i+1} &= \zeta^i + d\zeta, \\
\alpha^{i+1} &= \alpha^i + d\alpha,
\end{aligned}
$$

where $d\theta$ is given by Equation (18). However, due to the linear approximation, $({}^0\theta^{i+1}, \zeta^{i+1}, \alpha^{i+1})$ deviates slightly from the solution curve of Equations (16)-(17). So, the second step iterates the Newton algorithm to correct the predicted value of θ^{i+1}:

$${}^{k+1}\theta^{i+1} = {}^k\theta^{i+1} + \begin{bmatrix} J_1 \\ J_2 \end{bmatrix}^\dagger V({}^k\theta^{i+1}),$$

where V is a 6-D vector that represents an instantaneous motion from the current position and orientation to the desired (X, R). Once ${}^k\theta^{i+1}$ has converged, the limit is set as θ^{i+1} and the integration algorithm returns to the first step. The algorithm stops when ζ^{goal} and α^{goal} are reached. It is guaranteed to succeed whenever the forward kinematic map in Equations (16)-(17) are non-singular over the path from $(\zeta^{\text{init}}, \alpha^{\text{init}})$ to $(\zeta^{\text{goal}}, \alpha^{\text{goal}})$. Often such a path can be chosen as the line segment between $(\zeta^{\text{init}}, \alpha^{\text{init}})$ and $(\zeta^{\text{goal}}, \alpha^{\text{goal}})$ as non-canonical models differ only slightly from the canonical model. If the path contains a singularity of the forward kinematic map, our algorithm will not converge. When this happens, we choose a different path and run our algorithm again until a solution is found, or a designated number of iterations is reached. In the latter case, our algorithm reports failure.

Figure 18 shows two conformations sampled for the canonical 8-dof fragment and the corresponding two conformations obtained for a non-canonical model. Similarly, Figure 19
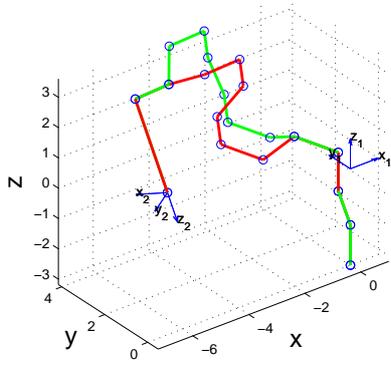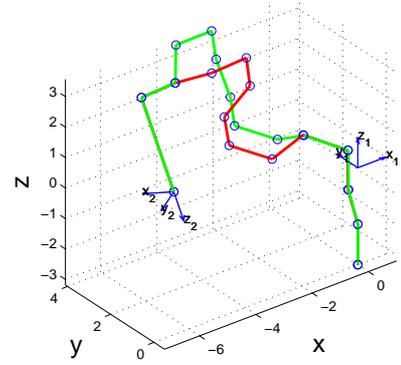
Figure 18: (a): Two conformations sampled for the canonical 8-dof fragment; (b): The corresponding two conformations obtained for the non-canonical fragment with $\zeta = 1.5\pi/180$ and $\alpha = 19\pi/180$.
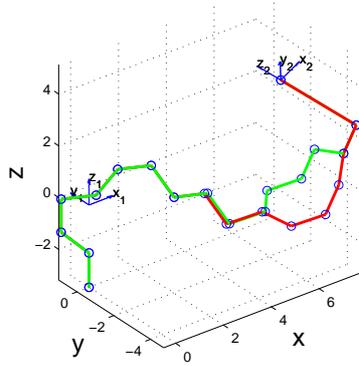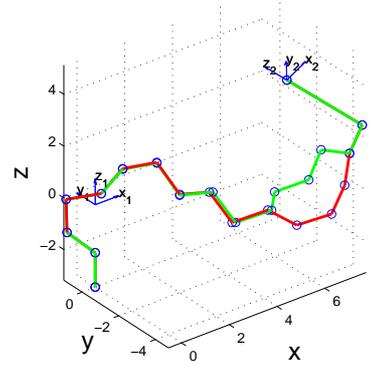
Figure 19: (a): Two conformations sampled for the canonical 10-dof fragment; (b): The corresponding two conformations obtained for the non-canonical fragment with $\zeta = 1.5\pi/180$ and $\alpha = 19\pi/180$.

shows two sampled conformations of a 10-dof fragment. In both examples, $\zeta = 1.5\pi/180$ and $\alpha = 19\pi/180$.

This homotopy algorithm can easily be extended to handle the case where all the $(\zeta_i^{\text{goal}}, \alpha_i^{\text{goal}})$ are distinct, as well as to the case where bond lengths and bond angles take residue-dependent values along the backbone fragment. Then the parameter vector and consequently the Jacobian matrices $\mathcal{V}_1$ and $\mathcal{V}_2$ have greater dimensionalities.

7 Conclusion

This paper presents a new approach to study the global structure of the IK map of a 6-dof fragment of protein and to sample conformations of a $2n$-dof fragment ($n > 3$) with two fixed anchors. Starting with a slightly idealized model of a 6-dof fragment, we split the IK map into an inverse orientation map and an inverse position map. We analyzed the critical sets of both maps, which yield a decomposition of $SO(3)$ and $\mathbb{R}^3$ into open regions where the number of IK solutions is constant. This allowed us to prove the existence a 16-solution region in $\mathbb{R}^3 \times SO(3)$. For longer fragments, we proposed a recursive procedure to sample IK solutions. The feasible ranges of the torsional dofs are successively computed and sampled to generate conformations that cover well a fragment's self-motion manifold. Finally we proposed a general numerical homotopy algorithm to deform conformations of canonical fragments into conformations of non-canonical fragments with the same number of dofs.

Our current work is aimed at combining these algorithms with collision detection algorithms to study and sample the clash-free subset of the self-motion manifold of a given fragment, taking all atoms in both the backbone and the side-chains into account.

References

[1] J.W. Burdick. On the inverse kinematics of redundant manipulators: characterization of the self-motion manifolds. *IEEE Int. Conf. on Robotics and Automation* (ICRA), pages $264-270$, May 1989.

[2] A.A. Canutescu and R.L. Dunbrack Jr. Cyclic coordinate descent: A robotics algorithm for protein loop closure. *Protein Science*, 12:963$-$972, 2003.

[3] S. Boyd and L. Vandenberghe. *Convex Optimization*, Cambridge University Press, 2004

[4] J. Cortés, T. Siméon, M. Renaud-Siméon, and V. Tran. Geometric algorithms for the conformational analysis of long protein loops. *J. Computational Chemistry*, 25(7):956$-$967, 2004.

[5] E.A. Coutsias, C Seok, M.P. Jacobson, and K.A. Dill. A kinematic view of loop closure. *J. Computational Chemistry*, 25:510$-$528, 2004.

[6] T. E. Creighton. *Proteins : Structures and molecular properties.* W. H. Freeman and Company, New York, 2nd edition, 1993.

[7] M.P. doCarmo, *Differential geometry of curves and surfaces.* Prentice-Hall, 1976.

[8] L.R. Dodd, T.D. Boone, and D.N. Theodorou. A concerted rotation algorithm for atomistic Monte Carlo simulation of polymer melts and glasses. *Molec. Phys.*, 78:961−996 1993.

[9] L.T. Watson, S.C. Billups, and A.P. Morgan. HOMPACK: A suitable codes for globally covergent homotopy algorithms, *ACM Transaction on Mathematical Software*, 13, 281-310.

[10] A. Enosh, S.J. Fleishma, N. Ben-Tal and D. Halperin. Assigning transmembrane segments to helices in intermediate-resolution structures. 12th *Int. Conf. on Intelligent Systems for Molecular Biology* (ISMB), pages 122−129, 2004.

[11] V. Guillemin and A. Pollack. *Differential Topology.* Prentice-Hall, 1974.

[12] R. Kolodny, L. Guibas, M. Levitt, and P. Koehl. Inverse kinematics in biology: the protein loop closure problem. *Int. J. Robotics Research*, 24:151−163, 2005.

[13] D. Manocha and Y. Zhu. Kinematic manipulation molecular chains subject to rigid constraints. *Int. Conf. on Intelligent Systems in Molecular Biology* (ISMB), 2:285−293, 1994.

[14] D. Manocha, Y. Zhu, and W. Wright. Conformational analysis of molecular chains using nano-kinematics. *Comput. Appl. Biosci.*, 11(1):71−86, 1995.

[15] C. Mavroidis and B. Roth. Structural Parameters which reduce the number of manipulator configurations. *J. Mech. Design, Trans. ASME*, 116:3−10, 1994.

[16] J. Milnor. *Morse Theory.* Princeton University Press, 1963.

[17] R.M. Murray, Z.X. Li, and S.S. Sastry. *A mathematical introduction to robotic manipulation.* CRC Press, 1994.

[18] F.P. Preparata and M.I. Shamos. *Computational geometry.* Springer-Verlag, New York, 1985.

[19] M. Raghavan and B. Roth. Inverse kinematics of the general 6R manipulator and related linkages. *J. Mech. Design, Trans. ASME*, 115:502−508, 1993.

[20] A. Shehu, L.E. Kavraki, and C. Clementi. Modeling protein conformational ensembles: From missing loops to equilibrium fluctuations. *Proteins: Struture, Function, and Bioinformatics*, 65:164-179, 2006.

[21] R.Singh and B. Berger. ChainTweak: Sampling from the neighbourhood of a protein conformation. *Pacific Symp. on Biocomputation*, Hawaii, January 2005.

[22] J.C. Trinkle and R.J. Milgram. Complete path planning for closed kinematic chains with spherical joints. *Int. J. Robotics Research*, 21(9):773−789, 2002.

[23] H. van den Bedem, I. Lotan, J.C. Latombe and A. M. Deacon. Real-space protein-model completion: an inverse-kinematics approach. *Acta Crystallographica*, D61:2−13, 2005.

[24] W.J. Wedemeyer and H.A. Scheraga. Exact Analytical Loop Closure in Proteins Using Polynomial Equations. *J. Computational Chemistry*, 20:819−844, 1999.

[25] M. Zhang and L.E. Kavraki. A new method for fast and accurate derivation of molecular conformations. *J. Chem. Info. Comp. Sci.*, 42(1):64−70, 2002.

Appendix A : Computation of $J\hat{\rho}$ (Section 3.2)

The Lie algebra $so(3)$ can be identified with the space of 3×3 skew-symmetric matrices $\{B \in \mathbb{R}^{3\times 3} \mid B + B^T = 0\}$. So, a basis for this space is:

$$J_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \ J_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \ J_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We have:

$$\frac{dR_{\tau_i}}{d\tau_i} = J_z R_{\tau_i},$$

and for any $R \in SO(3)$:

$$\begin{aligned}
RJ_x R^{-1} &= R_{11}J_x + R_{21}J_y + R_{31}J_z, \\
RJ_y R^{-1} &= R_{12}J_x + R_{22}J_y + R_{32}J_z, \\
RJ_z R^{-1} &= R_{13}J_x + R_{23}J_y + R_{33}J_z,
\end{aligned}$$

where R_{ij} denotes the element of R at the ith row and j^{th} column. With respect to the basis $\{J_x, J_y, J_z\}$, $RJ_x R^{-1}$, $RJ_y R^{-1}$, and $RJ_z R^{-1}$ become Rx, Ry, and Rz, respectively. So:

$$[D_{\tau_1}\hat{\rho}, D_{\tau_2}\hat{\rho}, D_{\tau_3}\hat{\rho}]\hat{\rho}^{-1} = [J_z, R_{\tau_1}LJ_z(R_{\tau_1}L)^{-1}, R_{\tau_1}LR_{\tau_2}LJ_z(R_{\tau_1}LR_{\tau_2}L)^{-1}],$$

and with respect to $\{J_x, J_y, J_z\}$, $J\hat{\rho}$ is the following matrix:

$$J\hat{\rho} = [z, R_{\tau_1}Lz, R_{\tau_1}LR_{\tau_2}Lz].$$

Appendix B: Quartic Surface Q (Section 4.2)

Recall from Section 4.2 that $Q = \{q(u, w)\}$, where:

$$q(u, w) = \begin{bmatrix} q_x \\ q_y \\ q_z \end{bmatrix} = \begin{bmatrix} c_u + c_\gamma c_w - s_\alpha s_\gamma s_w \\ -s_\alpha s_u + s_\alpha c_\gamma s_w + s_\gamma c_w \\ c_\alpha(s_u + s_w) \end{bmatrix}.$$

To derive the equation of the surface Q, we first consider a simple surface $\{\hat{q}(u, w)\}$ given by

$$\hat{q}(u, w) = \begin{bmatrix} \hat{q}_x \\ \hat{q}_y \\ \hat{q}_z \end{bmatrix} = \begin{bmatrix} s_u + s_w \\ c_u \\ c_w \end{bmatrix}.$$

This surface satisfies the following equation and is therefore quartic:

$$h(\hat{q}) := \hat{q}_x^4 + \hat{q}_y^4 + \hat{q}_z^4 + 2\hat{q}_x^2(\hat{q}_y^2 + \hat{q}_z^2) - 2\hat{q}_y^2\hat{q}_z^2 = 0.$$

On the other hand, q and $\hat{q}$ are related by a linear transformation $A(\alpha, \gamma)$:

$$\hat{q} = A(\alpha, \gamma)q.$$

Thus, Q is also a quartic surface and satisfies:

$$h(Aq) = 0. \tag{19}$$

The Jacobian matrix of q is:

$$Jq \;\; = \;\; \begin{bmatrix} \dfrac{\partial q}{\partial u} & \dfrac{\partial q}{\partial w} \end{bmatrix} \;\; = \;\; \begin{bmatrix} -s_u & -s_\alpha s_\gamma c_w - c_\gamma s_w \\ -s_\alpha c_u & s_\alpha c_\gamma c_w - s_\gamma s_w \\ c_\alpha c_u & c_\alpha c_w \end{bmatrix}.$$

It loses rank when $\tan(u) = \tan(w) = s_\alpha(1 + c_\gamma)/s_\gamma$. Hence, the singular points of q are $\{(u_0, w_0), (u_0, w_0 + \pi), (u_0 + \pi, w_0), (u_0 + \pi, w_0 + \pi)\}$, and the critical points are:

$$\begin{aligned}
q_1 &= [\tfrac{s_\gamma c_\alpha^2(1+c_\gamma)}{\sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}}, \;\; \tfrac{s_\gamma^2 c_\alpha^2}{\sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}}, \;\; \tfrac{2s_\alpha c_\alpha(1+c_\gamma)}{\sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}}]^T, \\
q_2 &= [\tfrac{s_\gamma(1+s_\alpha^2) - s_\gamma c_\gamma c_\alpha^2}{\sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}}, \;\; -\sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}, \;\; 0]^T, \\
q_3 &= [\tfrac{-s_\gamma(1+s_\alpha^2) + s_\gamma c_\gamma c_\alpha^2}{\sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}}, \;\; \sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}, \;\; 0]^T, \\
q_4 &= [\tfrac{-s_\gamma c_\alpha^2(1+c_\gamma)}{\sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}}, \;\; \tfrac{-s_\gamma^2 c_\alpha^2}{\sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}}, \;\; \tfrac{-2s_\alpha c_\alpha(1+c_\gamma)}{\sqrt{s_\gamma^2 + s_\alpha^2(1+c_\gamma)^2}}]^T,
\end{aligned}$$

where $u_0 = w_0 = \mathrm{atan}(s_\alpha(1 + c_\gamma)/s_\gamma)$. All tangent spaces of Q are 2-D, except at the four critical points q_i, $i = 1, ..., 4$, where they are 1-D. One can further show that the following two line intervals:

$$\lambda q_1 + (1 - \lambda)q_4 \;\;, \;\; 0 \leq \lambda \leq 1, \tag{20}$$
$$\lambda q_2 + (1 - \lambda)q_3 \;\;, \;\; 0 \leq \lambda \leq 1, \tag{21}$$

form the self-intersection set of Q.

We pose $\xi = \tan(t/2)$. The inverse map $\hat{p}_M^{-1}(X)$ for a given $X \in \mathbb{R}^3$ can be computed by substituting $q = X - [c_t, s_t, 0]^T$, $c_t = 1 - \xi^2/1 + \xi^2$, and $s_t = 2\xi/1 + \xi^2$ in Equation (19). The result is an 8th-order polynomial of ξ, whose roots can be derived using the Sturm method [5, 24].

Appendix C: Intersection of Two Circles on S^2 (Section 6.1)

Let us represent two generic circles on S^2 by:

$$\{\mathrm{Rot}(\nu_1, \beta_1)X_1 \mid \beta_1 \in S^1\} \quad \text{and} \quad \{\mathrm{Rot}(\nu_2, \beta_2)X_2, \mid \beta_2 \in S^1\},$$

respectively, where $\nu_i \in S^2$, $i = 1, 2$, are two distinct unit vectors, $\mathrm{Rot}(\nu_i, \beta_i)$ is the matrix representing the rotation of angle β_i about ν_i, and $\mathrm{Rot}(\nu_i, \beta_i)X_i$ is the vector obtained by applying $\mathrm{Rot}(\nu_i, \beta_i)$ to $X_i \in S^2$.

At each intersection between these circles, we pose $X_3 = \mathrm{Rot}(\nu_1, \beta_1)X_1 = \mathrm{Rot}(\nu_2, \beta_2)X_2$ and we express X_3 as a linear combination of ν_1, ν_2, and $\nu_1 \times \nu_2$, i.e., $X_3 = a_1\nu_1 + a_2\nu_2 + a_3(\nu_1 \times \nu_2)$. Since any rotation around ν_i leaves ν_i unchanged, we have: $X_3^T \nu_i = X_i^T \nu_i$. Thus:

$$a_1 + a_2\nu_1^T\nu_2 = X_1^T\nu_1,$$
$$a_1\nu_1^T\nu_2 + a_2 = X_2^T\nu_2,$$

and:

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 & \nu_1^T\nu_2 \\ \nu_1^T\nu_2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} X_1^T\nu_1 \\ X_2^T\nu_2 \end{bmatrix}.$$

In addition, since X_3 is a unit vector, we have:

$$a_1^2 + 2a_1a_2\nu_1^T\nu_2 + a_2^2 + \|\nu_1 \times \nu_2\|^2 a_3^2 = 1.$$

So:

$$a_3^2 = \frac{1 - a_1^2 - 2a_1a_2\nu_1^T\nu_2 - a_2^2}{\|\nu_1 \times \nu_2\|^2}.$$

A solution exists whenever $a_1^2 + 2a_1a_2\nu_1^T\nu_2 + a_2^2 \leq 1$. For each value of X_3, the values of θ_i, $i = 1, 2$, are the solutions of:

$$(X_i - (X_i^T\nu_i)\nu_i)^T(X_3 - (X_3^T\nu_i)\nu_i) = \|X_i - (X_i^T\nu_i)\nu_i\| \|X_3 - (X_3^T\nu_i)\nu_i\| \cos\theta_i,$$
$$\|(X_i - (X_i^T\nu_i)\nu_i) \times (X_3 - (X_3^T\nu_i)\nu_i)\| = \|X_i - (X_i^T\nu_i)\nu_i\| \|X_3 - (X_3^T\nu_i)\nu_i\| \sin\theta_i.$$

Appendix D: Convex Hull of the Minkowski Sum of $k \geq 3$ Circles (Section 6.1)

Consider Equation (15) reproduced below:

$$(R_3 + R_{\tau_2}LR_5 \cdots + R_{\tau_2}L \cdots LR_{\tau_{n-1}}LR_{2n-1})v_2 = LR_{-\tau_1}(X - v_0) - LR_{-\tau_1}R_1v_2.$$

Our goal is to approximate the 3-D volume spanned by the Minkowski sum of the $n - 1 \geq 3$ circles in the left-hand side by its convex hull. More generally, at each recursion of the algorithm described in Section 6.1, we consider an equation in which the left-hand side is the Minkowski sum of $k \geq 3$ circles, which we denote here by $K_1, ..., K_k$, and the right-hand side is a circle C. The k circles $K_1, ..., K_k$, have the same radius ℓ_1, but generally lie in distinct planes. We want to eventually compute the portion of C that lies within the bounding volume V of the Minkowski sum. Because the minimal convex set that contains V is its convex hull $\mathcal{H}(V)$, our goal here is to compute such a convex hull.

Proposition 1 *The convex hull $\mathcal{H}(V)$ of the Minkowski sum of k circles $K_1, ..., K_k$ is the Minkowski sum of their corresponding disks $D_1, ..., D_k$, where:*

$$D_i := \{\tilde{x}_i x_i + \tilde{y}_i y_i \mid x_i^2 + y_i^2 \leq \ell_1^2\}$$

and $\tilde{x}_i$, $\tilde{y}_i$ are the two unit base vectors for the plane of K_i.

31

Proof: It is obvious that $V \subseteq \sum_i D_i$. Since D_i is a disk, hence a convex set, $\sum_i D_i$ is also convex. Thus, $\mathcal{H}(V) \subseteq \sum_i D_i$.

On the other hand, $\sum_i C_i \subseteq V$. Consider two vectors $v^1 = v_{1,1} + v_2 + \cdots v_k$ and $v^2 = v_{1,2} + v_2 + \cdots v_k$ such that where $v_{1,1}, v_{1,2} \in K_1$, and $v_j \in K_j$ for $j = 2, \cdots, k$. They are both in V. So, according to the definition of $\mathcal{H}(V)$, we have:

$$\lambda v^1 + (1 - \lambda)v^2 \in \mathcal{H}(V), \ \lambda \in [0, 1].$$

Simplifying the left-hand side of the above equation yields:

$$(\lambda v_{1,1} + (1 - \lambda)v_{1,2}) + v_2 + \cdots + v_k \in \mathcal{H}(V)$$

This shows that $D_1 + K_2 + \cdots + K_k \in \mathcal{H}(V)$. Next, we let $v_1 \in D_1$, $v_{2,1}, v_{2,2} \in K_2$, $v_j \in K_j$ for $j > 2$. Since:

$$\lambda(v_1 + v_{2,1} + v_3 + \cdots + v_k) + (1 - \lambda)(v_1 + v_{2,2} + v_3 + \cdots + v_k) \in \mathcal{H}(V),$$

we have:

$$D_1 + D_2 + K_3 + \cdots + K_k \in \mathcal{H}(V).$$

Continuing this process yields:

$$\sum_i D_i \subseteq \mathcal{H}(V).$$

So, $\mathcal{H}(V) = \sum_i D_i$. ∎