Inverse Kinematics of Discretely Actuated Hyper-Redundant Manipulators Using Workspace Densities

Imme Ebert-Uphoff * Gregory S. Chirikjian †
Department of Mechanical Engineering
Johns Hopkins University
Baltimore, MD 21218

Abstract

Hyper-redundant manipulators present an alternative to conventional 6 DOF manipulators for inspection, space, and medical applications. The additional degrees of freedom facilitate obstacle avoidance and allow tasks to be performed even if some of the actuators fail. In this paper we consider hyper-redundant manipulators that are actuated discretely, e.g. using two-state actuators or motors with finite resolution. The inverse kinematics problem for a discretely actuated manipulator is intrinsically different from the one for its continuously actuated counterpart. We present a framework for the discussion of the discretely actuated case and propose an algorithm for the inverse kinematics. The algorithm generates solutions in linear time with respect to the number of manipulator actuators, as opposed to the exponential time required by brute force search.

1 Introduction

Hyper-redundant architectures include cascades of modules such as Stewart Platforms, variable geometry trusses, or serial links. We call any of these types macroscopically-serial manipulators. One particular subclass of hyper-redundant manipulators are mechanisms consisting of serial links connected by active revolute joints. They have been studied extensively, e.g. in the early 70's a prototype called the Active Cord Mechanism was built in Japan by Shigeo Hirose, who based the design and motion planning procedures on studies on snakes [1]. More recent developments include a prototype of a Hyper-redundant Active Endoscope for use in remote and minimally invasive surgery [2], and a prototype with 12 degrees of freedom that is currently being tested at the Jet Propulsion Laboratory for inspection tasks on space stations [3]. A wide variety of algorithmical issues have been discussed for this kind of manipulator, see for example [4, 5] and references therein.

*Graduate Student
†Assistant Professor, Presidential Faculty Fellow

A second class of hyper-redundant manipulators uses variable geometry truss mechanisms (VGT's) stacked on top of one another. This design originated from space applications. Actuators were added to large static space structures so that they could be folded during transport and actively deployed in space. An advantage of this type of manipulator is that it offers excellent static stability. Different candidate topologies and their suitability for applications in space have been discussed extensively, e.g. [6, 7, 8].

One approach to the inverse kinematics of general macroscopically-serial manipulators is to prescribe their shape using backbone curves ([9, 10] and references therein).

In this paper we consider manipulators of general macroscopically-serial architecture that are actuated discretely. We consider discrete actuators because they offer high precision and reliability with low cost compared to actuators with continuous range-of-motion. For example, a continuous actuator can be turned into a very precise two-state actuator by replacing the position feedback by a mechanical constraint and overpowering it, to make sure that both states are reached under load. This is also an alternative for applications where position sensors and feedback are not available or are too expensive.

The inverse kinematics problem for a discretely actuated manipulator is intrinsically different from the one for its continuously actuated counterpart. Assuming continuous motion capabilities to determine ideal joint values and rounding these to the closest discrete joint values available can result in large errors. The reason for this is that the discrete joint values closest to the ideal joint values do not necessarily lead to the discrete configuration closest to the ideal configuration (since the forward kinematics are in general non-linear). In addition small inaccuracies close to the base can result in large displacements at the tip.

We present a framework for the discussion that considers only the discrete configurations attainable by the manipulator and develop an algorithm that solves the inverse kinematics problem in linear time with respect to the number of manipulator modules. The algorithm is based on determining and fixing a configuration for each of the modules sequentially starting...
Figure 1: A macroscopically-serial manipulator
at the base while maximizing at every step the redundancy remaining to reach the goal. A measure for the remaining redundancy is derived from characteristics of the workspace of the more distal platforms. The approach is illustrated for binary manipulators (using two-state actuators), as an example of a hyper-redundant manipulator with extremely low resolution in each actuator.

2 Concepts for Discrete Manipulators

This section provides concepts and definitions for manipulators whose joints have only discrete states. The discussion is restricted to manipulators of the macroscopically-serial type, i.e. a serial collection of modules where each module is mounted on top of the previous one. Modules are numbered 1 to \( B \), from the base to the end of the manipulator, as shown in Figure 1. Closed loops may exist in each module, but macroscopic loops are not permitted.

2.1 Point Density and Intermediate Workspaces

The workspace of a continuous range-of-motion manipulator is often described by its boundary and the ease with which points in the interior can be reached is described using dexterity measures. A discrete manipulator can only reach a finite number of points. Therefore not only the boundary of the workspace is important but also the distribution of the points inside this boundary.

For discrete manipulators one has to distinguish carefully between redundant architecture and redundancy of motion. In the context of discretely actuated manipulators redundancy of motion only has meaning when an error threshold is specified. That is, given an error radius \( r \), a discrete manipulator is redundant for a specified position and orientation (frame), if there is more than one configuration within the distance \( r \) of the frame. The distance is measured according to a distance function that compares two frames with each other, e.g. such as in [11].

Depending on whether the manipulator under consideration is planar or spatial, and whether orientation is considered, the manipulator workspace \( W \) is a subset of \( \mathbb{R}^2 \), \( \mathbb{R}^2 \times SO(2) \), \( \mathbb{R}^3 \) or \( \mathbb{R}^3 \times SO(3) \).

From now on we assume that the workspace \( W \) is divided into blocks (pixels) of equal size. The distribution of the points is described as follows: The point density \( \rho \) assigns each block of workspace \( W \) the number of points within the block that are reachable by the discrete manipulator, normalized by the volume of the block:

\[
\rho(\text{block}) = \frac{\# \text{ reachable points in block}}{\text{unit volume/area}}
\]

The point density is a probabilistic measure of the positional accuracy of the end-effector in a considered area of the workspace. The higher the density in the neighborhood of a point, the more accurately we expect to be able to reach the point.

The \( i \)th intermediate workspace, \( W_i \), of a macroscopically-serial manipulator composed of \( B \) modules is the workspace of the manipulator segment from module \( i + 1 \) to the end-effector. End-effector position vectors in the \( i \)th intermediate workspace are written in a coordinate frame attached to the base of the module \( i \). To visualize this, imagine that the manipulator arm is cut between module \( i \) and module \( i + 1 \). The two resulting segments are considered as manipulators on their own: the lower segment (modules 1 to \( i \)), and the upper segment (modules \( i + 1 \) to \( B \)) with its base in the separating plane. The workspace generated by the upper manipulator segment is the \( i \)th intermediate workspace of the whole structure.

Each module has a frame attached to its top. The frames are numbered such that frame \( F_i \) is on top of module \( i \), and frame \( F_0 \) is the frame at the manipulator base. The number of discrete configurations of module \( i \) is denoted \( K_i \). The \( K_i \) discrete configurations of module \( i \) (\( i = 1, \ldots, B \)) are represented by the configuration set:

\[
C_i = \{(\mathbf{R}_1^{(i)}, \mathbf{b}_1^{(i)}), (\mathbf{R}_2^{(i)}, \mathbf{b}_2^{(i)}), \ldots, (\mathbf{R}_{K_i}^{(i)}, \mathbf{b}_{K_i}^{(i)})\},
\]

where \( \mathbf{R}_j^{(i)} \in SO(N) \) are rotation matrices and \( \mathbf{b}_j^{(i)} \in \mathbb{R}^N \) are translation vectors for \( j = 1, \ldots, K_i \). These pairs describe all possible relative orientations and positions of frame \( i \) with respect to frame \( i - 1 \).

2.2 The Configuration tree

In this subsection we present a computational structure for manipulators with discrete joint resolu-
tion: the configuration tree. In later sections we will explain how it can be used to solve the inverse kinematics for manipulators with low joint resolution, in particular for binary manipulators.

As described in the preceding subsection we consider macroscopically-serial manipulators. The modules are numbered, starting at the base with module 1, and proceeding up to the most distal module with the number B. The configuration tree (CT) of such a manipulator is a tree defined by the following conventions:

- Starting from the root of the tree with module 1, each level represents one consecutive module of the manipulator.
- At level i each node has exactly one child for each configuration attainable by module i. Each of these \( K_i \) branches has a geometric description of the module configuration assigned to it: \( (R_{ij}^{(1)}, b_{ij}^{(1)}) \in C_i \).

As a result the nodes of the tree hold the following information:

- The nodes of level i enumerate all possible configurations of the partial manipulator given by module 1 to i.
- The leaves of the tree, i.e. the nodes of the last level, enumerate all configurations the whole manipulator can attain.
- For any node on level i the relative orientation and position of frame \( F_i \) with respect to the fixed base frame \( F_0 \) can be determined in the following way: we walk from the root of the tree to the desired node, consecutively applying all geometric relationships assigned to traversed links.

The inverse kinematics can now be implemented as a search in the tree. For the decision process the densities of intermediate workspaces are used. Note that the configuration tree is never generated completely, since the computational complexity is exponential in the number of modules \( \mathcal{O}(\prod_{i=1}^{B} K_i) \). Instead paths in the tree are only generated as necessary.

3 Inverse Kinematics

In this section we explain our approach for the inverse kinematics problem, avoiding the exponential complexity required by brute force computation. Given a target frame, we try to reach it with the end-effector of the manipulator.

The discussion consists of two parts. In the first part a measure is developed that determines how well the goal can be reached if modules close to the base are fixed in a particular configuration. The second part presents algorithms that use this measure to solve the inverse kinematics problem.

3.1 A measure for reachability

3.1.1 Generation of intermediate workspaces

As a prerequisite for the measure of reachability we need a description of intermediate workspaces that includes their internal point density.

Although every discrete manipulator has only a finite number of states, it is impossible to store (or calculate) the end-effector positions for all states of a discrete hyper-redundant manipulator due to exponential complexity \( \mathcal{O}(\prod_{i=1}^{B} K_i) \). On the other hand, if the workspace is discretized in a finite number of pixels of equal size, it is possible to quickly approximate and store the number of points contained in each of these pixels, which in turn can be converted to the point density.

The algorithm used in this paper to approximate the point densities of all intermediate workspaces is called the Mapping Algorithm and is described in [12] in much detail. Also included in that paper
are numerical examples and an error bound on the approximation. 1 The following gives a brief overview of the algorithm.

The Mapping Algorithm determines intermediate workspaces starting at the end-effector and ending at the base. At each step we climb down one module, maintaining an approximation of the intermediate workspace corresponding to the segment of all modules above the current one. The point density of each intermediate workspace is described by an \(N\)-dimensional array of integers representing the number of points in each pixel of the discretized workspace, where \(N\) is the dimension of the workspace. Combined with administrative information, i.e. the pixel dimensions and the coordinates of a reference point of the workspace, this array represents the density distribution in the workspace.

Each iteration of the algorithm is based on the observation that workspace \(W_{i-1}(\text{workspace of modules } i \text{ to } B)\) is connected to workspace \(W_i (\text{workspace of modules } i + 1 \text{ to } B)\) through the set \(C_i\) of all possible configurations of module \(i\). To generate the point density of workspace \(W_{i-1}^T\), the algorithm interprets the elements of \(C_{B-i}\) as homogeneous transformations and applies them to workspace \(W_i^T\). Superposition of the results yields workspace \(W_{i-1}^T\). Note, that the complexity of each homogeneous transformation depends only on the number of pixels chosen for the discretization of workspace \(W_i^T\) and can be implemented very efficiently as an identical operation on all indices of the array. The generation of all intermediate workspaces is performed offline and requires time linear in the number of modules of the manipulator (\(O(B)\)), linear memory, and quadratic but small error.

In more general terms, this transformation can be reformulated as a convolution of functions on the Euclidean group, \(SE(N)\). Two real-valued functions, \(f_1(H)\) and \(f_2(H)\), where \(H \in SE(N)\), are used to describe the distribution of configurations \(C_{B-i}\) and the point density of intermediate workspace \(W_i^T\), respectively. The convolution of the two functions \(^2\) results in the density distribution \(f_3(H)\) of the intermediate workspace \(W_i^T\):

\[
f_3(H) = (f_1 * f_2)(H) = \int_{SE(N)} f_1(\mathcal{H}) f_2(\mathcal{H}^{-1}H) d\mu(\mathcal{H}),
\]

where \(d\mu(\mathcal{H})\) is a volume element in \(SE(N)\). An algorithm for workspace generation based on this equation as well as mathematical and physical interpretation of above equation are described in [13]. For identical

1 Although the algorithm is described in [12] only for the case of binary actuation, it can be generalized to other discrete actuation by a simple change, namely the numbering convention for the configurations of the modules has to be adjusted.

2 \(f_1\) and \(f_2\) have to be nonnegative and bounded everywhere, and decrease rapidly enough to zero such that the convolution integral converges.

![Figure 3: Manipulator divided into two segments](image)

modules the workspace generation is achieved in order \(O(\log B)\) time.

### 3.1.2 Derivation of Reachability Measure

The goal of this subsection is to develop a measure that predicts how well a goal specified by position vector \(b(\text{ee})\) and rotation matrix \(R(\text{ee})\) can be reached by the end-effector if modules close to the base are fixed in a particular configuration.

The manipulator is divided into a lower segment (modules 1 to \(i\), and an upper segment (modules \(i + 1\) to \(B\), as described in section 2.1 and shown in Figure 3. To describe their configurations we use the following notation:

- \((R(L), b(L))\): orientation and position of frame \(F_i\) with respect to frame \(F_0\) (lower segment)
- \((R(U), b(U))\): orientation and position of frame \(F_B\) with respect to frame \(F_i\) (upper segment)
- \((R(\text{ee}), b(\text{ee}))\): target orientation and position of frame \(F_B\) with respect to frame \(F_0\) (whole manipulator)

Composing the frame transformations \((R(L), b(L))\) and \((R(U), b(U))\), one finds that

\[
\begin{align*}
(R(\text{ee}), b(\text{ee})) &= (R(L) R(U), b(L) + R(L)b(U))
\end{align*}
\]

Consequently, the ideal configuration the upper module has to attain for any particular configuration of the lower segment is described by

\[
\begin{align*}
R(U) &= \left(R(L)\right)^T R(\text{ee}) \\
b(U) &= \left(R(L)\right)^T (b(\text{ee}) - b(L))
\end{align*}
\]

Recall, that the \(i\)th intermediate workspace of the manipulator is the workspace of the upper segment when considered as a manipulator on its own. The point density of this workspace is available through the methods mentioned in 3.1.1. We evaluate the point density at the pixel of the ideal orientation and position of the upper segment, \((R(U), b(U))\). The higher the point density is around \((R(U), b(U))\), the more ac-
curately we expect to reach \((R^{(e)}, b^{(e)})\) with the end-effector.

**Evaluation of reachability measure:** In summary the measure on the configuration \((R^{(U)}, b^{(U)})\) of the lower module is implemented as:

1. Calculate the homogeneous transform \((R^{(U)}, b^{(U)})\) for the upper segment to reach the target.
2. Evaluate the density of the pixel of the i-th intermediate workspace in which \((R^{(U)}, b^{(U)})\) lies to determine how well it can be reached.

Note, that the evaluation of the reachability measure in the examples for this paper is only based on the position \(b^{(U)}\) and neglects the information of the orientation \(R^{(U)}\). Consequently, the algorithms for the inverse kinematics that are derived for the general case (including orientation) in this section, are only simulated in positional coordinates.

### 3.2 Algorithms for Inverse Kinematics

Our discussion of the inverse kinematics problem consists of two parts. The first part shows how to resolve the geometric constraints arising from the task of reaching a target location and orientation as closely as possible. An algorithm that finds a solution to these constraints is presented in the first subsection. The second part of this subsection discusses strategies to resolve the remaining redundancy to avoid abrupt transitions between configurations.

#### 3.2.1 Satisfying the geometric constraints

The criterion for reachability established in the previous subsection can be used in a simple algorithm that determines a solution for the given target position and orientation. The algorithm starts with the first module and considers the finite set of its possible configurations. For each configuration the reachability measure is evaluated as described in 3.1 and the algorithm picks the one with the maximal value. The module is fixed in this configuration and the algorithm continues with the next module. The following describes the algorithm in terms of the configuration tree:

**Path of Probability Algorithm (POP)**

1. Input: target location and orientation. Start with the root of the tree.
2. For all children of the current node: evaluate the reachability criterion and choose the branch that maximizes the reachability criterion. Take the child as current node.
3. If the current node does not have any children: Output configuration. DONE. Otherwise: goto 2.

The reachability criterion can be evaluated in constant time. Considering the maximal number of configurations per module as a constant, the total order of this algorithm is of the number of modules, \(\mathcal{O}(B)\), and is therefore also linear in the number of actuators.

#### 3.2.2 Resolving the remaining redundancy

The algorithm of the previous section (POP) provides a solution close to the target position and orientation and hence solves the inverse kinematics problem. In practice, however, we are not only interested in finding the solution to one frame but to a set of frames or a trajectory, expecting some kind of smooth and repeatable motion. This is not guaranteed by the preceding algorithm. The problem is that since the redundancy in the interior of the workspace is not resolved, the algorithm arbitrarily picks one out of a number of possible solutions.

Our approach to resolve the redundancy is to maximize a secondary criterion in the algorithm. The secondary criterion prescribes characteristics of the motion in some way while it resolves the remaining redundancy. Examples of secondary criteria are:

- The weighted number of modules that remain in the same state as in the previous configuration, with a higher weight close to the base.
- The weighted number of modules that are fully contracted (i.e. binary state 0), with a higher weight close to the base.

While the first criterion above compares two configurations with each other, (the previous one with the one under consideration), the second is independent of previous motion (history-independent). The second criterion above generates behavior similar to a long narrow spring: the manipulator expands only as far as necessary. In general history-independent criteria are preferred because they guarantee higher repeatability.

**Improved Path of Probability (IPOP)**

1. Input: target location and orientation. Start with the root of the tree.
2. For all children of the current node: evaluate the reachability criterion and one of the secondary criteria. Choose one branch based on a combination of the two criteria.
3. If the current node does not have any children: Output configuration. DONE. Otherwise: goto 2.

Our implementation of step 2 considers all configurations for which the primary criterion is above a certain threshold (to bound the error). Out of these we choose the one that satisfies the secondary criterion the best.


4 Simulations of Binary Manipulator Inverse Kinematics

The inverse kinematics algorithms are applied to a planar binary manipulator that consists of $B = 16$ variable geometry truss modules. Each module has three independent actuators which can only attain the lengths 0.2 and 0.25, while the base is fixed at width 0.2. This manipulator can reach $2^{316} = 2^{48} \approx 2.8 \times 10^{14}$ points in the workspace. The state of the manipulator can be described as a bit sequence, using one bit for each actuator. State 0 indicates that the actuator is contracted, state 1 indicates that it is extended. The actuators are numbered from right to left, from the most distal module to the module at the base, as shown in Figure 4 for a manipulator consisting of only four modules.

The simulations are performed on a Pentium PC (60 MHz) using NextStep. The time required to calculate the inverse kinematics for a single position ranges from 50 to 120 msec, depending on the location in the workspace.

Figures 5(a-c) show results of the inverse kinematics using three different algorithms. The task is to reach three points with the end-effector, (3.0, 0.7), (0.55, 2.9), (0.4, 2.7), in this order, starting from an upright contracted position (bit sequence 0). In the background of each figure the point density of the manipulator workspace is shown in a logarithmic gray scale. Point densities above a certain threshold are shown in identical, dark gray. The three target points are marked by black crosses. The bit sequences corresponding to all generated configurations are listed in Figure 6, together with resulting end-effector position, orientation and error. Each bit sequence is written in the octal system, i.e. each digit corresponds to the configuration of exactly one module.

In Figure 5(a) the POP algorithm is applied, Figure 5(bc) apply the IPOP algorithm. Figure 5(b) uses similarity to previous configurations as the secondary criterion while Figure 5(c) uses maximal contraction as the secondary criterion starting from the base.

Note that the configurations chosen to reach the point furthest to the right, (3.0, 0.7), look very much alike in all figures. The reason is that the target point is close to the workspace boundary, where the density is low, thus restricting the shape of modules close to the base to be unique, as can be seen by comparing the most significant bits of the corresponding bit sequences in Figure 6. In contrast, the points (0.55, 2.9) and (0.4, 2.7) are in an area with very high point density and a significant redundancy is left after the geometric constraints are satisfied. The results from the three different figures for these two points give a good example of how the different algorithms resolve the redundancy. The POP algorithm resolves this redun-
dancy randomly, i.e., a configuration is picked arbitrarily from those within an error threshold. Hence the solutions in Figure 5(a) for the two points differ substantially. In Figure 5(b) modules are as similar as possible to the previous configuration with a higher weight given to modules close to the base. As a result the configurations for the upper two points are similar to each other and, as much as possible, to the configuration for the right point. Using maximal contraction as the secondary criterion, as shown in Figure 5(c), the solutions for the upper points are also similar to each other although the solution for each point is independent from the others.

5 Conclusion

This paper discusses an efficient way to solve the inverse kinematics of discretely actuated manipulators with many degrees of freedom.

A measure is developed which can test partial configurations for their suitability to reach a specific goal. Based on this measure an algorithm (IPOP) for the inverse kinematics is proposed that generates solutions satisfying the geometric constraints of the task. The remaining redundancy is resolved by additional constraints imposed on the algorithm. This improved form of the algorithm (IPOP) avoids abrupt transitions between configurations. Different constraints are discussed and demonstrated by simulations.

References


<table>
<thead>
<tr>
<th>target (x, y)</th>
<th>(3.0, 0.7)</th>
<th>(0.55, 2.9)</th>
<th>(0.42, 2.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>high 24 bits</td>
<td>46.46.67</td>
<td>0.00.05.24</td>
<td>0.00.05.24</td>
</tr>
<tr>
<td>low 24 bits</td>
<td>0.00.12</td>
<td>1.51.65.21</td>
<td>0.02.22.12</td>
</tr>
<tr>
<td>POP</td>
<td>x = 3.000</td>
<td>x = 0.550</td>
<td>x = 0.397</td>
</tr>
<tr>
<td></td>
<td>y = 0.700</td>
<td>y = 2.898</td>
<td>y = 2.900</td>
</tr>
<tr>
<td></td>
<td>θ = -1.621</td>
<td>θ = 0.039</td>
<td>θ = 0.000</td>
</tr>
<tr>
<td></td>
<td>e = 0.002</td>
<td>e = 0.002</td>
<td>e = 0.003</td>
</tr>
<tr>
<td>IPOP - previous</td>
<td>46.46.00.66</td>
<td>46.46.31.31</td>
<td>46.46.31.31</td>
</tr>
<tr>
<td></td>
<td>77.16.16.76</td>
<td>27.23.61.64</td>
<td>37.16.77.60</td>
</tr>
<tr>
<td>IPOP - contracted</td>
<td>x = 2.998</td>
<td>x = 0.550</td>
<td>x = 0.401</td>
</tr>
<tr>
<td></td>
<td>y = 0.699</td>
<td>y = 2.899</td>
<td>y = 2.899</td>
</tr>
<tr>
<td></td>
<td>θ = -1.845</td>
<td>θ = -0.264</td>
<td>θ = 0.001</td>
</tr>
<tr>
<td></td>
<td>e = 0.092</td>
<td>e = 0.001</td>
<td>e = 0.001</td>
</tr>
<tr>
<td></td>
<td>46.46.00.66</td>
<td>00.00.00.00</td>
<td>00.00.00.00</td>
</tr>
<tr>
<td></td>
<td>77.16.16.76</td>
<td>62.06.12.73</td>
<td>40.25.70.60</td>
</tr>
</tbody>
</table>

Figure 6: Results for different end-effector coordinates from different algorithms.


