

On the Disentanglement of Tube Inequalities in Concentric Tube Continuum Robots

Reinhard M. Grassmann^{*,1}, Anastasiia Senyk^{*,1,2}, and Jessica Burgner-Kahrs¹, *Senior Member, IEEE*

Abstract—Concentric tube continuum robots utilize nested tubes, which are subject to a set of inequalities. Current approaches to account for inequalities rely on branching methods such as if-else statements. It can introduce discontinuities, may result in a complicated decision tree, has a high wall-clock time, and cannot be vectorized. This affects the behavior and result of downstream methods in control, learning, workspace estimation, and path planning, among others.

In this paper, we investigate a mapping to mitigate branching methods. We derive a lower triangular transformation matrix to disentangle the inequalities and provide proof for the unique existence. It transforms the interdependent inequalities into independent box constraints. Further investigations are made for sampling, control, and workspace estimation. Approaches utilizing the proposed mapping are at least 14 times faster (up to 176 times faster), generate always valid joint configurations, are more interpretable, and are easier to extend.

I. INTRODUCTION

A concentric tube continuum robot (CTCR) is composed of several nested concentric tubes being pre-curved, and super-elastic [1], [2]. The inherent compliant and flexible nested tubes are translated and rotated with respect to each other to change the manipulator’s shape. Different tube materials can achieve this elastic interaction between tubes. Such robots typically comprise two or three metal alloy tubes, usually nitinol [3]. CTCR robot’s distinct feature lies in the nonlinear dynamics resulting from the elastic interplay between these tubes. On top of this, from a mechanical and modeling point of view, the use of nested tubes imposes tube length constraints formulated as a set of two inequalities. Unfortunately, little attention is paid to mitigating the effects of these inequalities.

Consider a CTCR with two tubes being subject to the two inequalities $0 \geq \beta_1 \geq \beta_2$ and $0 \leq L_1 + \beta_1 \leq L_2 + \beta_2$. As illustrated in Fig. 1, the translational joint space \mathcal{B} forms a parallelogram. Both inequalities span the edges of the translational joint space \mathcal{B} . From this observation, we can use rotation, scaling, and shearing to obtain

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}}_{\text{shearing}} \underbrace{\begin{bmatrix} L_1 & 0 \\ 0 & L_2 - L_1 \end{bmatrix}}_{\text{scaling}} \underbrace{\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}}_{\text{rotation}} \underbrace{\begin{bmatrix} \beta_{1,\mathcal{U}} \\ \beta_{2,\mathcal{U}} \end{bmatrix}}_{\in \mathcal{U}^2} = \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}}_{\in \mathcal{B}}$$

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [RGPIN-2019-04846] as well as the Canada Foundation for Innovation and Ontario Research Fund [Project #40110].

¹Continuum Robotics Laboratory, Department of Mathematical and Computational Sciences, University of Toronto, Mississauga, ON L5L 1C6, Canada. reinhard.grassmann@utoronto.ca

²Faculty of Applied Science, Ukrainian Catholic University, Lviv, Ukraine.

*Equal contribution.

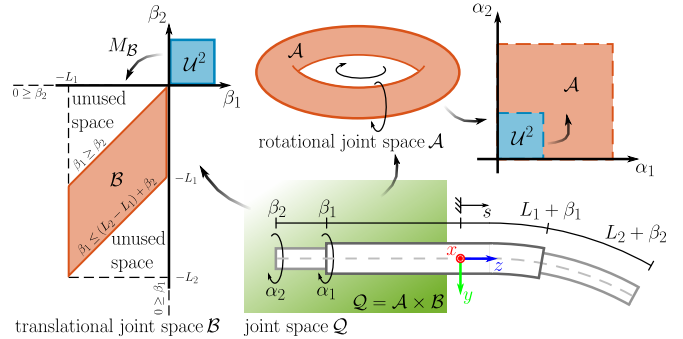


Fig. 1. Characteristics of the joint space of a 4 dof CTCR with two nested tubes. While tube rotations α_i form a torus, therefore, can be represented as square, the tube translations β_i result in a parallelogram. This parallelogram is bounded by the inequalities. Considering this geometric insight, it follows that an affine transformation M_B can be used to map a square \mathcal{U}^2 to the parallelogram \mathcal{B} . The geometric insight extends to N tubes, and it should be clear that the transformed space, i.e., \mathcal{U}^N for N tubes, is more desirable.

We introduced this type of transformation in [4]. For learning the kinematics of a CTCR task, we show in [4] that the accuracy and convergence are significantly improved by using this simple yet effective transformation to decorrelate the joint space. As of now, this approach has been used in publications [4], [5], [6] on machine learning. Yet, we are confident that disentanglement will also be useful for other research fields in the continuum robotics research community. Furthermore, looking at applications beyond CTCRs, the disentanglement of inequalities is also useful for continuum robot types with variable lengths and cranes with prismatic joints. For instance, a tendon-driven continuum robot’s variable segments can be achieved using tubes [7].

More importantly, one or both inequalities are usually neglected for approaches purely evaluated in simulation, e.g., [8], [9], [10]. Observing the proliferation of prototypes [11] poses a risk for translating the results on physical hardware due to the need for considering both inequalities. Furthermore, if both inequalities are considered, methods in the literature rely on branching methods. Branching methods such as if-statements can be implemented straightforwardly. However, they can introduce discontinuities, have high wall-clock time, and are poorly vectorizable. Consequently, methods and approaches using branching methods may not be suitable for controlling, learning, optimization, workspace estimation, and path planning. Especially computationally expensive applications or applications considering small time steps would suffer from a high wall-clock time.

In this paper, we expand on our previously proposed affine transformation matrix [4], [5] by providing a proof

for its uniqueness, discussing its theoretical foundation, and showing its effectiveness in applications beyond learning, i.e., sampling, workspace estimation for concentric tube continuum robots, and control synthesis.

II. DISENTANGLEMENT OF BETA

In this section, we recap the affine transformation matrix to disentangle the translational joint space as illustrated in Fig. 1. While this transformation is proposed in our previous work [4], [5], we present more theoretical insights and identify patterns by explicitly generalizing to N tubes. On top of this, we prove the uniqueness of the affine transformation and show that only one such mapping exists.

A. Affine Transformation Matrix

Due to the use of nested tubes, both inequalities given by

$$0 \geq \beta_1 \geq \beta_2 \geq \dots \geq \beta_N \quad \text{and} \quad (1)$$

$$0 \leq L_1 + \beta_1 \leq L_2 + \beta_2 \leq \dots \leq L_N + \beta_N, \quad (2)$$

have to be satisfied by each translation β_i . Here, the constant L_i is the length of the i^{th} tube. The interdependencies caused by (1) and (2) result from the fact that none of the distal ends of inner tubes should be inserted inside the respective outer tube. This can be derived both from a theoretical point of view [12] and from a practical point of view, e.g., to avoid play between adjacent tubes, which can cause hysteresis [13]. Furthermore, (1) can prevent collision between carriages, while (2) can prevent collision between the outer tube and gripper or attached sensors [5].

To avoid the evaluation of the inequalities (1) and (2), the translations β_i can be transformed such that the new joint space representation $\beta_{i,\mathcal{U}}$ are decorrelated. For the sake of compactness, we summarize all N translations β_i and $\beta_{i,\mathcal{U}}$ in a column matrix β and $\beta_{\mathcal{U}}$, respectively. For the transformation $\beta = M_{\mathcal{B}}\beta_{\mathcal{U}}$, the mapping matrix given by

$$M_{\mathcal{B}} = \begin{bmatrix} -L_1 & 0 & 0 & \dots & 0 \\ -L_1 & L_1 - L_2 & 0 & \dots & 0 \\ -L_1 & L_1 - L_2 & L_2 - L_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -L_1 & L_1 - L_2 & L_2 - L_3 & \dots & L_{N-1} - L_N \end{bmatrix} \quad (3)$$

is defined for N tubes and follows the used convention in [5] and shown in Fig. 1, i.e., most inner tube $i = N$ and outermost tube $i = 1$. The determinant of (3) is

$$\det M_{\mathcal{B}} = (-1)^N L_1 \prod_{i=2}^N (L_i - L_{i-1}). \quad (4)$$

The factors of the product are the eigenvalues, and the sign of the determinant changes depending on if N is odd or even. More importantly, since $0 < L_{i-1} < L_i$, the inverse of (3)

always exists. The inverse of the lower triangular matrix is

$$M_{\mathcal{B}}^{-1} = \begin{bmatrix} \frac{1}{-L_1} & 0 & 0 & \dots & 0 \\ \frac{1}{L_2 - L_1} & \frac{1}{L_1 - L_2} & 0 & \dots & 0 \\ 0 & \frac{1}{L_3 - L_2} & \frac{1}{L_2 - L_3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{1}{L_{N-1} - L_N} \end{bmatrix} \quad (5)$$

being a matrix with nonzero values in the lower diagonal and main diagonal. By computing $M_{\mathcal{B}}^{-1}\beta$, all entries of $\beta_{\mathcal{U}}$ are in the interval $[0, 1]$.

To enforce $\beta_{i,\mathcal{U}} \in [-1, 1]$, the matrix (3) is scaled and shifted. Removing the homogeneous extension leads to

$$\beta = \left[\frac{1}{2}M_{\mathcal{B}} \quad \frac{1}{2}M_{\mathcal{B}} \cdot \mathbf{1}_{N \times 1} \right] \beta_{\mathcal{U}}. \quad (6)$$

Inverting the homogeneous extension of (6) leads to

$$\beta_{\mathcal{U}} = [2M_{\mathcal{B}}^{-1} \quad -\mathbf{1}_{N \times 1}] \beta \quad (7)$$

without the homogeneous extension, where all $\beta_{i,\mathcal{U}}$ are automatically scaled between -1 and 1 as well as unit-less and orthogonal. The previously interdependent translational parameters β_i subject to (1) and (2) are now disentangled. In fact, the inequalities become independent N box constraints, i.e., $-1 \leq \beta_{i,\mathcal{U}} \leq 1$.

B. Considering Minimal Length

It is worth noting that L_i is not necessarily the physical length of the tube. We follow an approach that is similar to the configuration approach in [14] and capture additional mechanical constraints resulting from the actuation unit. This gives the advantage that $\beta_i \in [\beta_{i,\min}, 0]$, where $\beta_{i,\min} = -L_i$. In order to consider a safety margin for sensors [5], the minimum displacement denoted by $\beta_{i,\min}$ can be further restricted by considering a margin $L_{i,\text{margin}}$. We can define

$$\beta_{i,\min} = - \left(L_i - \sum_{k=i}^N L_{k,\text{margin}} \right) = -L_i^* \quad (8)$$

as the minimum displacement, where the restriction is applied recursively. Afterwards, L_i in (3) and (5) are substituted by $L_i^* = L_i - \sum_{k=i}^N L_{k,\text{margin}}$ to obtain $M_{\mathcal{B}}^*$ and $M_{\mathcal{B}^*}^{-1}$, where the star notation indicates the consideration of L_i^* .

C. Uniqueness of the Low Triangular Transformation Matrix

Now, we illustrate the derivation of (3). The main idea is to first independently select a value for β_k of the k^{th} tube. Consequently, values for the adjoint tubes, i.e., $(k-1)^{\text{th}}$ tube and $(k+1)^{\text{th}}$ tube, depend on the chosen β_k . This step is repeated until the most inner and outermost tubes are reached and all values are chosen. Consequently, for a given set of N different tubes, it is possible to find N different starting points for this approach and eventually N different affine

transformation matrices. However, only one of them satisfies both inequalities (1) and (2), which is (3).

Starting with the outermost tube, we see the pattern

$$\beta_1 = -L_1\beta_{1,\mathcal{U}} \quad (9)$$

$$\beta_2 = \beta_1 - (L_2 - L_1)\beta_{2,\mathcal{U}} = \beta_1 + (L_1 - L_2)\beta_{2,\mathcal{U}}$$

⋮

$$\beta_i = \beta_{i-1} + (L_{i-1} - L_i)\beta_{i,\mathcal{U}} \quad (10)$$

emerging, where $\beta_{i,\mathcal{U}} \in \mathcal{U}[0, 1]$ is a uniform distributed random variable. Introducing $\beta_0 = 0$ and $L_0 = 0$, the first sampling step (9) is considered in pattern (10), which can be used as base case for an induction proof. Furthermore, looking at the lower bound, i.e., $\beta_{i,\mathcal{U}} \geq 0$, we can obtain a short hand for the inequality (1) being $\beta_i \leq \beta_{i-1}$. In addition, considering the upper bound, i.e., $\beta_{i,\mathcal{U}} \leq 1$, the pattern (10) simplifies to $\beta_i \leq \beta_{i-1} + (L_{i-1} - L_i)$. After rearranging leading to $\beta_i + L_i \leq \beta_{i-1} + L_{i-1}$, the other inequality (2) can be distilled. The result shows that (10) considers both inequalities, including the case for $i = 0$. Using induction, the critical fact that the pattern (10) is always negative, i.e., $\beta_i \leq 0$, is reinforced as the base case (9) is negative and $L_{i-1} - L_i \leq 0$.

Forward substitution is used to obtain $M_{\mathcal{B}}$. From (10), it is clear that this leads to a lower triangular matrix, where the entries of a column are either zero or the coefficient of $\beta_{i,\mathcal{U}}$ in (10). More importantly, the subsequently derived $M_{\mathcal{B}}$ given by (3) takes both inequalities (1) and (2) into account.

Now, we look into an alternative pattern. For this purpose, one can start with the most inner tube, i.e., $\beta_N = -L_N\beta_{N,\mathcal{U}}$, as the base case. The emerging pattern is $\beta_i = \beta_{i+1} + (L_{i+1} - L_i)\beta_{i,\mathcal{U}}$. This pattern leads to an upper triangular matrix. Descending to $i = 0$, the pattern results in $\beta_1 = -L_1\beta_{0,\mathcal{U}}$, where $\beta_{0,\mathcal{U}}$ is undefined. More critical, it is not ensured that β_i is always negative. To show this, we attempt a proof by contradiction. Assuming the emerging pattern is positive, i.e., $0 \leq \beta_{i+1} + (L_{i+1} - L_i)\beta_{i,\mathcal{U}}$, for the $i = N - 1$ case, we get $0 \leq \beta_N + (L_N - L_{N-1})\beta_{N,\mathcal{U}}$. Rearranging and substituting the base case, this yields $\beta_{N,\mathcal{U}} \leq (L_N - L_{N-1})/L_N\beta_{N-1,\mathcal{U}}$. Hence, the attempt fails, and a condition between the random variables $\beta_{N,\mathcal{U}}$ and $\beta_{N-1,\mathcal{U}}$ exists, where β_N can be positive. Note that, for (10), this would lead to a contradiction since the right side of (10) is always negative. Therefore, starting with the inner tube does not lead to a valid mapping respecting (1) and (2).

Considering a starting point between the most inner and outermost tube, i.e., $\beta_k = -L_k\beta_{k,\mathcal{U}}$ for the k^{th} tube with $1 < k < N$, this would lead to two patterns – the ascending pattern (10) and the faulty descending pattern. Hence, any of the $N - 2$ remaining matrices cannot guarantee that (1) holds. In conclusion, $M_{\mathcal{B}}$ is the only linear mapping that utilizes the stated idea of successive selecting β_k and considers both inequalities.

III. APPLICATION TO JOINT SPACE SAMPLING

In this section, we compare and contrast four rejection sampling methods with a sampling method utilizing the

mapping $M_{\mathcal{B}}$. For evaluation, we use a Monte Carlo method to determine the distribution and aggregate each rejection sampling method's average time and success rate. A unit distribution is used of each β_i , where the three tubes lengths are set to 100 mm, 150 mm, and 200 mm.

A. Rejection Sampling via Branching

Rejection sampling is the most common approach to sample β_i subject to (1) and (2). All values of the k^{th} sample are sampled independently from a uniform distribution, i.e.,

$$\beta^{(k)} \leftarrow \text{diag}(-L_i)\beta_{\mathcal{U}}^{(k)} \quad (11)$$

where $\text{diag}(-L_i)$ is short for $\text{diag}(-L_1, -L_2, -L_3)$. Afterward, both inequalities are checked for the k^{th} sample, and, if necessary, a resampling approach is applied. This step is repeated until (1) and (2) hold.

To achieve a desired sample size, we consider four methods with different resampling approaches. If the k^{th} sample $\beta^{(k)}$ does not satisfy both inequalities (1) and (2), we

- (a) re-sample all values,
- (b) re-sample while keeping $\beta_1 = -L_1\beta_{1,\mathcal{U}}$,
- (c) re-sample while keeping $\beta_2 = -L_2\beta_{2,\mathcal{U}}$, or
- (d) re-sample while keeping $\beta_3 = -L_3\beta_{3,\mathcal{U}}$.

The first approach, i.e., (a), is the most common variant in the literature. The other methods can be considered as a special case, where a specific joint value is preset.

B. Direct Sampling with Affine Transformation Matrix

Another approach is to map a sampled point $\beta_{\mathcal{U}}$ from a uniform cube \mathcal{U}^3 using $M_{\mathcal{B}}$. As proven in Sec. II-C, all samples fulfill (1) and (2). For the k^{th} sample, we can utilize

$$\beta^{(k)} = M_{\mathcal{B}}\beta_{\mathcal{U}}^{(k)} \quad (12)$$

to sample a valid joint values $\beta_i^{(k)}$ using (3).

C. Evaluation and Results

Figure 2 shows the distribution of each sampling method for 5000 samples. The sample population is collected by running each method five times for 1000 samples. Table I lists the wall-clock time and success rate. While a Matlab script on a consumer PC is used to generate the data, the relative increase of the wall-clock time is relevant for an implementation-independent comparison.

The four sampling methods (a)-(d), branching – implemented via if-else statement – and resampling – implemented as a while loop – are used, causing larger wall-clock time than the direct sampling method, which is at least two magnitudes faster. On the contrary, our sampling method using $M_{\mathcal{B}}$ is a branchless approach and offers three main advantages. First, the success rate is 100%. Second, the execution is fast as it only relies on matrix multiplication with complexity of $O(N(N + 1)/2)$. Third, the wall-clock time can be further improved by vectorizing (12), where all samples can be computed in one step, which is impossible with other sampling methods.

TABLE I

WALL-CLOCK TIME AND SUCCESS RATE FOR DIFFERENT SAMPLING METHODS. THE STATED FACTOR IS THE NORMALIZED WALL-CLOCK TIME W.R.T. THE DIRECT SAMPLING METHOD. BASED ON A MAXIMAL RESAMPLING ATTEMPT OF 1000, THE FAIL RATE INDICATES HOW OFTEN A SAMPLING METHOD FAILS TO CONVERGE TO A VALID SAMPLE W.R.T ALL SUCCESSFUL SAMPLING ATTEMPTS. NOTE THAT THE COMPLEMENT OF THE SUCCESS RATE IS THE RATE WITH WHICH THE SAMPLED VALUES ARE NOT VALID AND MUST BE RESAMPLED.

method	time in ms	factor	fail rate	success rate
(a)	58.8 ± 3.6	14.1	N/A	8.4 %
(b)	80.4 ± 3.7	19.3	0 %	8.3 %
(c)	189.1 ± 10.5	45.3	0.5 %	3.2 %
(d)	614.5 ± 81.3	147.3	5.5 %	1 %
M_B	4.2 ± 0.3	1	N/A	100 %
M_B (vectorized)	0.3 ± 0.3	0.08	N/A	100 %

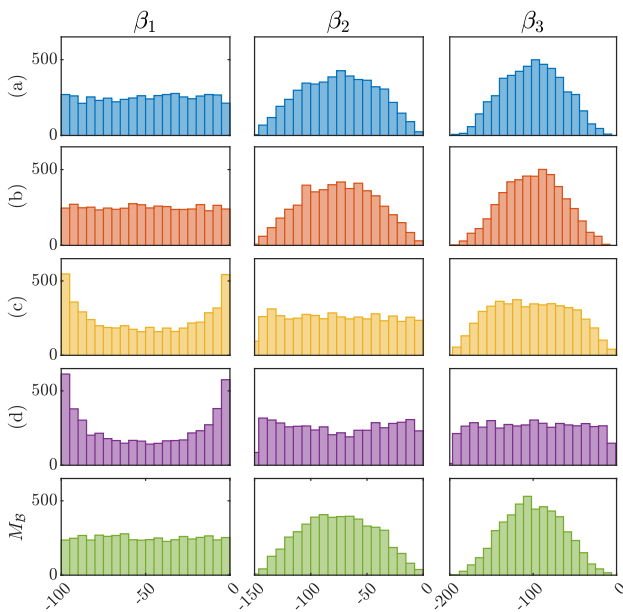


Fig. 2. Distributions for different sampling method. The columns refer to the respective sampled joint values β_i , whereas the rows refer to the used sampling method. The rejection sampling methods (a) to (d) are described in Sec. III-A. The direct sampling using M_B is described in Sec. III-B.

For the simple and most commonly used approach, the success rate of the resampling approach (a) can be computed directly. Using (4), success rate is equivalent to

$$\frac{|\det M_B|}{\prod_{i=1}^N L_i} = \prod_{i=2}^N \frac{L_i - L_{i-1}}{L_i} = \prod_{i=2}^N \left(1 - \frac{L_{i-1}}{L_i}\right) < 1, \quad (13)$$

being the ratio of the volume. The theoretical value of 8.3 % is consistent with the empirical result 8.4 % listed in Table I. From (13), each factor depends on the ratio of adjacent tubes and is smaller than one. These facts reveal two characteristics. First, the similar the lengths of the adjacent tubes are, the lower the success rate. Second, the more nested tubes, the lower the success rate. As a consequence, the lower (13), the larger the wall-clock time.

Note the relation between (9) and the sampling method (11) with (b). From that, it is no surprise that the unbiased sampling method (a), sampling method (b), and the sam-

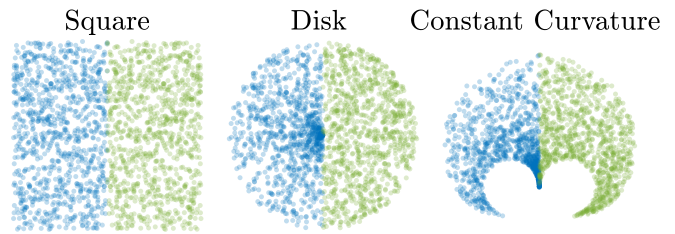


Fig. 3. Affecting of distribution on two-dimensional workspace of toy examples. Two variables are randomly drawn from uniform distributions and used afterward to generate the workspace for a square, a disk, and a single-segment constant curvature continuum robot. Each workspace is divided into two areas – (blue) vanilla approach and (green) proposed approach. The blue sampled points are linearly transformed, whereas, for the green sampled points, with the exception of the squared workspace, the additional transformation (14) is used for the translational variable. As can be seen from the realized distribution, the green sampled points yield a more homogeneous distribution.

pling method using M_B converges to similar distributions. Looking into M_B and the pattern (10) leading to (3), we can see that β_i is the weighted sum of i number of different independent scaled uniform distributions. Therefore, the distribution of each β_i relates to an Irwin-Hall distribution – also known as the uniform sum distribution – explaining the characteristics of distributed joint values for the three sampling methods. This effect is clearly noticeable with a higher sample size, i.e., in [5], the distributions shows a clear uniform, triangular, or bell-shaped distribution. By contrast, the sampling methods (c) and (d) can result in a U-shaped distribution. Note that the higher the value for the maximum resampling attempt, see Table I, the more pronounced the U-shaped distribution for β_2 using (d). These U-shaped characteristics can be seen in [15] too. Moreover, the wall-clock time of (b) and (c) is significantly higher, while the success rate is low. Further note that we added an upper bound on the resampling attempts, see Table I, such that (a), (b), and (c) terminate after a finite number of steps, otherwise the wall-clock time would have increased further. This directly translates to downstream methods that use a resampling approach with preset joint values.

IV. APPLICATION TO WORKSPACE ESTIMATION

We will discuss an observation before applying our sampling method to workspace estimation. The observation relates to the sampling of two variables to cover an area. After presenting a solution to the observation, we briefly state a method to exploit rotational symmetries. Finally, this section ends with a comparison of workspace estimation given different sampling methods.

A. Translation and the Square Root Transformation

We start with toy examples as motivation. To cover an area, two random variables are required. For the simple case of a rectangle, sampling both independent dimensions with uniform distribution results in expected distributed points in the area. In contrast, when sampling a disk, the points are undesirably distributed, see Fig. 3. Moreover, a planar workspace of a continuum robot with constant curvature

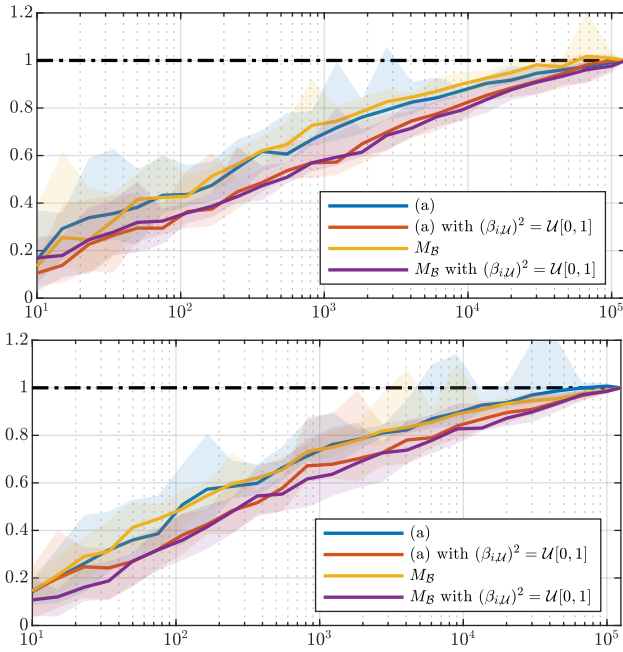


Fig. 4. Convergence of workspace estimation in percent over the number of samples. The top and bottom plot show the workspace estimation for a CTCR with geometrical parameters stated in [5] and [15], respectively. The solid lines are the median values being limited by the minimum and maximum estimates. Ten different permutations of the ordered dataset are used to compute the median, minimum, and maximum area. Note that a new sampled point between two existing points at the boundary can reduce the computed area by the triangular area spanned by those three points and, therefore, the dashed 100% mark can be exceeded. Further note that the workspace estimation using M_B has a faster overall time of convergence.

kinematics and variable length is closely related to a circular area. With no surprise, it shows a similar effect, see Fig. 3.

For the disk, one approach is to derive the Jacobian determinant of the transformation between the Cartesian and polar coordinate [16], [17]. This leads to a simple solution for the disk, where the radius is determined by sampling its square. Adapting this solution for $\beta_{i,U}$ related to the tube length, this leads to

$$(\beta_{i,U})^2 = \mathcal{U}[0, 1]. \quad (14)$$

Afterward, the mapping (3) can be used. Note that, for the disk, the angle is independent of the length variable. However, this is not the case for the constant curvature. We leave the proper sampling of the curvature for future work.

B. Workspace Estimation

To evaluate our proposed approach, we determine the workspace of two CTCRs. We compare the sampling method stated in (11) with the resampling approach (a), and (12). Both sampling methods are used with and without (14). For the former, $\beta_U^{(k)} = \sqrt{\mathcal{U}^N[0, 1]}$ is used in place of $\beta_U^{(k)} = \mathcal{U}^N[0, 1]$. To create each dataset, 125 000 tip positions are sampled for each sampling method and CTCR. To compute the tip position, we use a static model [12] with Poisson's ratio of 0.3, Young's Modulus of 50 GPa, and the geometrical parameters are stated in [15] and [5].

The sampled tip positions are mapped onto a plane via as the workspace of the CTCR exhibits rotational symmetry.

TABLE II
AREA AND CLOSENESS OF SAMPLED WORKSPACES.

method	$\sqrt{\mathcal{U}^N}$	closeness			CTCR ²
		mean	median ¹	area	
(a)	✗	0.056 ± 0.237	0.052	0.0175	[5]
(a)	✓	0.038 ± 0.195	0.032	0.0144	[5]
M_B	✗	0.056 ± 0.237	0.053	0.0171	[5]
M_B	✓	0.040 ± 0.200	0.034	0.0134	[5]
(a)	✗	0.049 ± 0.221	0.045	0.0131	[15]
(a)	✓	0.037 ± 0.193	0.032	0.0115	[15]
M_B	✗	0.049 ± 0.221	0.045	0.0131	[15]
M_B	✓	0.037 ± 0.192	0.032	0.0103	[15]

¹Median of the medians is computed instead of the median directly since $S = 125\,000$ samples require to compare $S(S-1)/2$ distances, which is too big – 58.2 GB – to fit in the memory of our machine.

²Reference to the parameters of the CTCR. For [15], the *Robot 1* defined in Table I is used.

This can be achieved either by plotting the distance between a point and the symmetry axis against the symmetry axis, or by rotation each point into a plane.¹ Afterwards, a boundary (*boundary* in MATLAB with shrink factor set to one) of the set of points is found and the area (*polyarea* in MATLAB) is computed. Figure 4 shows the results of the workspace estimation. Table II lists the estimated area as well as the closeness. The closeness is measured by the Euclidean distance between all projected tip positions.

While the absolute value reported in Table II is less descriptive, the relative differences between sampling methods with and without (14) show that the mean and median distance are smaller and less spread according to the standard deviation. From this, it can be concluded that using (14) can distributed the sampled tip position more evenly. However, at the expense of a smaller estimated workspace. From Fig. 4, the sampling method without (14) converges faster to the final area at the expense of less homogeneous distributed samples in the area. Moreover, as expected, (11) with (a) re-sampling approach and (12) perform similarly. The exception here is the workspace estimation, where (12) with (14) tends to generate a smaller area.

V. APPLICATION TO CONTROL

In this section, we will incorporate M_B to a control application. This is motivated by the fact that a closed-loop system with a simple control for β_i can output values that do not satisfy (1) and (2). This is possible even if all inputs are valid desired $\beta_{i,d}$. For the sake of simplicity, a first-order proportional delay element (PT₁) model and a PI controller with $K_{P,i}$ and $K_{I,i}$ gains for the i^{th} tube are used for the closed-loop control. This leads to a simple linear and time-invariant system. The block diagram is shown in Fig. 5.

A. Vanilla Approach

The vanilla approach treats each β_i independently. Therefore, the block diagram looks similar to the one shown in

¹By rotating all tip positions onto one plane, we like to call this the “Lantern method” as it reminds us of the opening and unfolding of a paper lantern. This method [18], [19] is a simple yet useful simplification.

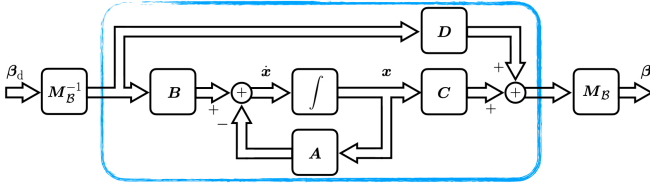


Fig. 5. Block diagram of the closed-loop system in state-space. The system encased in blue represents a simple PT_1 model of the simplified motor dynamics controlled by a PI controller. This system must account for the two interdependent inequalities (1) and (2). By using M_B , they become independent box constraints that are easier to handle.

Fig. 5, where the additional transformation using M_B is set to the identity matrix. The matrices of the state equation are

$$A = \begin{bmatrix} \text{diag}(-K_{P,i} - 1) & \text{diag}(K_{I,i}) \\ -\mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \quad (15)$$

and $B = [\text{diag}(K_{P,i}), \mathbf{I}_{3 \times 3}]^T$. The output equations are given by $C = [\mathbf{I}_{3 \times 3}, \mathbf{0}_{3 \times 3}]$ and $D = \mathbf{0}_{3 \times 3}$.

B. Transforming the State Space Representation

A rather novel approach utilizing (3) is to transform the state space vector as shown in Figure 5. Using block diagram algebra, the state equation changes to

$$\hat{A} = [\text{diag}(M_B, M_B)] A [\text{diag}(M_B^{-1}, M_B^{-1})], \quad (16)$$

where $\text{diag}(M_B, M_B)$ and $\text{diag}(M_B^{-1}, M_B^{-1})$ are block diagonal matrices whose diagonal contains blocks of M_B and M_B^{-1} , respectively. From (16), we get $M_B \text{diag}(-K_{P,i} - 1) M_B^{-1}$ resulting in

$$\begin{bmatrix} -K_{P,1} - 1 & 0 & 0 \\ K_{P,2} - K_{P,1} & -K_{P,2} - 1 & 0 \\ K_{P,2} - K_{P,1} & K_{P,3} - K_{P,2} & -K_{P,3} - 1 \end{bmatrix}, \quad (17)$$

whereas the upper right block, i.e., $M_B \text{diag}(K_{I,i}) M_B^{-1}$, is

$$\begin{bmatrix} K_{I,1} & 0 & 0 \\ K_{I,1} - K_{I,2} & K_{I,2} & 0 \\ K_{I,1} - K_{I,2} & K_{I,2} - K_{I,3} & K_{I,3} \end{bmatrix}. \quad (18)$$

The matrices $\hat{C} = C$ and $\hat{D} = D$ stay unchanged, while

$$\hat{B} = \begin{bmatrix} M_B \text{diag}(K_{P,i}) M_B^{-1} \\ \mathbf{I}_{3 \times 3} \end{bmatrix} \quad (19)$$

is the transformed B , where $M_B \text{diag}(K_{P,i}) M_B^{-1}$ has a similar appearance to (17), only it negative and without ones.

C. Advantages

While inequalities (1) and (2) can be incorporated and considered in an optimal control scheme, using M_B , the inequalities simplify to box constraints. Hence, the controller to be synthesized can be more computationally efficient compared to an optimal control scheme. In other words, transforming the state space representation leads to a system, which can be treated as a linear time-invariant system. Therefore, a larger variety of methods are at the disposal of the user. Furthermore, nonlinear elements such as saturation and anti-windup elements can be integrated with ease.

Another benefit is the analysis of such transformed system, which can be straightforward. For instance, for (16) of the transformed system, we can see that the constraints

$$K_{P,i} \leq K_{P,i+1} \quad \text{and} \quad K_{I,i+1} \leq K_{I,i} \quad (20)$$

for (17) and (18), respectively, are necessary otherwise the system becomes a non-minimum phase system. Therefore, an undesirable set of controller gains can make the closed-loop system unstable. This insight gives rise to possible challenges in control synthesis for CTCRs, when using a branching method similar to Sec. III-A.

VI. DISCUSSION

Using M_B has been shown beneficial for machine learning application [4], [5], [6]. In this paper, we touch on the application in control, workspace estimation, and, in a more general sense, in sampling. While the results are preliminary, they show promising future direction utilizing the advantage of M_B being the transformation of (1) and (2) to N box constraints. This opens the possibility to use branchless approaches offering several advantages. Among others, it reduces the time spent on branch prediction, resulting in faster execution. This is particularly beneficial for real-time applications that require quick responses, such as control systems. Additionally, in computational hungry applications like tube design or online learning, where reducing time is crucial to ensure efficient resource utilization. Another benefit of utilizing M_B is that it gives the possibility to shape distributions to any desired distributions using the inverse transform sampling.

We acknowledge that other representations and notations exist. For example, in [20], [21], the translation is defined by the segment length of overlapping tubes, i.e., $l_i = \beta_i + L_i - (\beta_{i-1} + L_{i-1}) \geq 0$. This definition considers only the manipulator side, i.e., (2), which has its merits for algorithm in simulation. However, the algorithm applied to real hardware must account for (1), which reflects the actuation side. Moreover, l_i either needs to be measured or related to β_i , which does not solve the inequalities in the first place. Furthermore, indexing starts either with the outermost tube, e.g., [22], [23], [5], or with the most inner tube, e.g., [24], [4]. We advocate using the former, showing advantages in formalizing patterns, sequence, and proofs as used in Sec. II-C. In any case, caution is advised when comparing formulae and algorithms using different notations. For example, the upper triangular matrix in [4] is the low triangular matrix (3) and vice versa.

VII. CONCLUSIONS

In this paper, we investigate the affine transformation matrix M_B to mitigate the use of branching methods. Its derivation and uniqueness are shown. Use of this transformation has ripple effects on downstream approaches previously using sampling or branching. We explored the application to sampling, workspace estimation, and control showing advantages in success rates, lowering the computational load, and interpretability of used methods.

REFERENCES

- [1] P. Sears and P. Dupont, "A steerable needle technology using curved concentric tubes," in IEEE/RSJ international conference on intelligent robots and systems, 2006, pp. 2850–2856.
- [2] R. J. Webster, A. M. Okamura, and N. J. Cowan, "Toward active cannulas: Miniature snake-like surgical robots," in IEEE/RSJ international conference on intelligent robots and systems, 2006, pp. 2857–2863.
- [3] C. J. Nwafor, C. Girerd, G. J. Laurent, T. K. Morimoto, and K. Rabenoroso, "Design and fabrication of concentric tube robots: A survey," IEEE Transactions on Robotics, 2023.
- [4] R. Grassmann and J. Burgner-Kahrs, "On the merits of joint space and orientation representations in learning the forward kinematics in se(3)," in Robotics: Science and Systems, 2019.
- [5] R. M. Grassmann, R. Z. Chen, N. Liang, and J. Burgner-Kahrs, "Dataset and benchmark for learning the kinematics of concentric tube continuum robots," in IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2022, pp. 9550–9557.
- [6] K. Iyengar, S. H. Sadati, C. Bergeles, S. Spurgeon, and D. Stoyanov, "Sim2real transfer of reinforcement learning for concentric tube robots," IEEE Robotics and Automation Letters, 2023.
- [7] T.-D. Nguyen and J. Burgner-Kahrs, "A tendon-driven continuum robot with extensible sections," in IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2015, pp. 2130–2135.
- [8] M. Boushaki, C. Liu, and P. Poignet, "Task-space position control of concentric-tube robot with inaccurate kinematics using approximate jacobian," in International Conference on Robotics and Automation. IEEE, 2014, pp. 5877–5882.
- [9] C. Fellmann, D. Kashi, and J. Burgner-Kahrs, "Evaluation of input devices for teleoperation of concentric tube continuum robots for surgical tasks," in Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling. International Society for Optics and Photonics, 2015, pp. 411 – 419.
- [10] S. Sabetian, T. Looi, E. Diller, and J. Drake, "Self-collision detection and avoidance for dual-arm concentric tube robots," Robotics and Automation Letters, pp. 1–8, 2019.
- [11] R. M. Grassmann, S. Lilge, P. Le, and J. Burgner-Kahrs, "Cter prototype development: An obstacle in the research community?" in Robotics: Science and Systems – Workshop on Robotics Retrospectives, 2020.
- [12] D. C. Rucker, B. A. Jones, and R. J. Webster III, "A geometrically exact model for externally loaded concentric-tube continuum robots," IEEE transactions on robotics, vol. 26, no. 5, pp. 769–780, 2010.
- [13] C. Greiner-Petter and T. Sattel, "On the influence of pseudoelastic material behaviour in planar shape-memory tubular continuum structures," Smart Materials and Structures, vol. 26, no. 12, p. 125024, 2017.
- [14] T. Lozano-Perez, "Spatial planning: A configuration space approach," IEEE Transactions on Computers, vol. 32, no. 2, pp. 108–120, 1983.
- [15] J. Burgner-Kahrs, H. B. Gilbert, J. Granna, P. J. Swaney, and R. J. Webster, "Workspace characterization for concentric tube continuum robots," in IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2014, pp. 1269–1275.
- [16] S. R. Dunbar, "The average distance between points in geometric figures," The College Mathematics Journal, vol. 28, no. 3, pp. 187–197, 1997.
- [17] B. Eisenberg and R. Sullivan, "Crofton's differential equation," The American Mathematical Monthly, vol. 107, no. 2, pp. 129–139, 2000.
- [18] C. Bergeles, A. Gosline, N. Vasilyev, P. Codd, P. del Nido, and P. Dupont, "Concentric tube robot design and optimization based on task and anatomical constraints," Transactions on Robotics, vol. 31, no. 1, pp. 67–84, 2015.
- [19] G. Noh, S. Yoon, S. Yoon, K. Kim, W. Lee, S. Kang, and D. Lee, "Expeditious design optimization of a concentric tube robot with a heat-shrink plastic tube," in International Conference on Intelligent Robots and Systems. IEEE/RSJ, 2016, pp. 3671–3676.
- [20] L. Lyons, R. Webster, and R. Alterovitz, "Motion planning for active cannulas," in International Conference on Intelligent Robots and Systems. IEEE/RSJ, 2009, pp. 801–806.
- [21] —, "Planning active cannula configurations through tubular anatomy," in International Conference on Robotics and Automation. IEEE, 2010, pp. 2082–2087.
- [22] R. Webster, J. Romano, and N. Cowan, "Mechanics of precurved-tube continuum robots," Transactions on Robotics, vol. 25, no. 1, pp. 67–78, 2008.
- [23] Y. Baran, K. Rabenoroso, G. Laurent, P. Rougeot, N. Andreff, and B. Tamadazte, "Preliminary results on oct-based position control of a concentric tube robot," in International Conference on Intelligent Robots and Systems. IEEE/RSJ, 2017, pp. 3000–3005.
- [24] X. Yang, S. Song, L. Liu, T. Yan, and M. Meng, "Design and optimization of concentric tube robots based on surgical tasks, anatomical constraints and follow-the-leader deployment," Access, vol. 7, pp. 173 612–173 625, 2019.