Rapid Path-Planning for Unstable Concentric Tube Robot Guidance

Konrad Leibrandt, Christos Bergeles, Member, IEEE, and Guang-Zhong Yang, Fellow, IEEE

Abstract—The complex, non-intuitive kinematics of concentric tube robots can make their telemanipulation challenging. Collaborative control schemes that guide the operating clinician via repulsive and attractive force feedback based on intraoperative path-planning can simplify this task. Computationally efficient algorithms, however, are required to perform rapid path-planning and to solve the inverse kinematics of the robot at interactive rates. Until now, ensuring stable and collision-free robot configurations required long periods of pre-computation to establish kinematic look-up tables. This paper presents a high-performance robot kinematics software architecture, which is used together with a multi-node computational framework to rapidly calculate dense path-plans for safe telemanipulation of unstable concentric tube robots. The proposed software architecture enables on-the-fly incremental inverse-kinematics estimation at interactive rates, and is tailored to modern computing architectures with efficient multi-core CPUs. The effectiveness of the architecture is quantified with computational-complexity metrics, and in a clinically demanding simulation inspired from neurosurgery for hydrocephalus treatment. By achieving real-time path-planning we can generate active constraints on-the-fly and support the operator in faster and more reliable execution of telemanipulation tasks.

I. INTRODUCTION

Accessing pathologies through natural orifices or single ports for minimally invasive surgery (MIS) is a challenging task well suited for continuum robots as their shape flexibility allows them to controllably conform to the traversed anatomy [3]. Concentric tube robots (CTRs) represent an example of continuum robots, and entail the translation and rotation of pre-curved super-elastic tubes that are concentrically arranged [4]. The pre-curved tubes mutually interact to create a final robot shape, allowing CTR to exhibit full tip-pose control. As the number of tubes of the robot increases, its redundant configuration space can be exploited to simultaneously achieve shape control and tip manipulation. By controlling the number and mechanical properties of tubes, CTRs can be deployed for a variety of surgical applications, such as haemorrhage evacuation [5], transurethral prostatectomy [6], tissue approximation for heart surgery [7] and bronchoscopic biopsy [8]. Computational design of CTR can deliver optimal robot architectures for different surgical tasks [7], [9]. The computational requirements for safe inverse kinematics increase with the number of robot tubes that these tailored robots possess, supporting the need for quick and efficient solvers.

Detailed models based on tube mechanics have been developed for the forward and inverse kinematics of CTRs. The most advanced models account for the torsional wind-up along the length of each tube, as well as the instability of configurations (equilibrium conformations) arising from torsion [10], [11]. Energy accumulated through the torsional wind-up of the tubes can rapidly release in certain tube configurations and cause uncontrollable motion [11], [12]. Therefore, unstable configurations need to be actively avoided. However, their detection incurs additional computational cost and constraints on the inverse kinematics.

Real-time, safe and effective path-planning for custom-designed CTRs has several benefits. First, it empowers the operator with a reliable telemanipulation approach, ensuring avoidance of anatomical collisions and dangerous unstable configurations [1], [7], [13], [14], [15]. Second, if real-time performance is guaranteed, it is possible to custom-make concentric tubes for specific patients using rapid workspace optimisation with minimal time delay. Especially in critical cases such as ruptured aneurysms [5], a rapid plan-to-deployment period can significantly reduce intra-operative complication and potential morbidity. Finally, real-time path-
planning schemes can be coupled with active constraints (AC)/virtual fixtures (VF) [16] to form cooperative robotic telemanipulation controllers wherein the operator is gently steered towards the estimated safe paths. In this shared-control approach, the complexity of inverse kinematics is invisible to the operator, allowing seamless avoidance of collisions and unstable configurations while retaining overall control over the approach path to the anatomical targets.

Fast solvers for robot inverse kinematics are a primary building block towards assisted telemanipulation of CTRs. To this end, researchers have achieved real-time kinematics performance by creating look-up tables and employing root-finding approaches to estimate the joint values for a desired pose of the tip [10]. Others have investigated Jacobian-based approximations [17] to speed-up computations. Pre-computation of dense path-plans within the robot workspace has been proposed by [18], while sparse path-plans and random trees have been proposed in [14], [19].

Most commonly, the solvers do not consider stability and prune the robot-design space to avoid unstable robots. This approach, however, limits the combinations of tubes that can be used, despite [7] having demonstrated that unstable CTRs may well be the only robots that can perform certain interventions. In addition, pre-computation of path-plans assumes a static anatomy and complicates the intraoperative adaptation of the said plans when anatomy is dynamic. Finally, pre-computation for achieving interactive-rate kinematics solution adds many hours of overhead in computational requirements - even when instability is not considered. Only faster approaches to inverse kinematics estimation can deliver the fundamental building blocks for safe and effective CTR telemanipulation in dynamic environments. Our proposed approach is well suited to tackle these challenges.

Our work expands on the algorithms of [1], [2]. The proposed software architecture takes advantage of distributed computing resources and multi-core CPUs to provide globally dense workspaces (robot task space) within minutes. By creating memory-optimal data structures to represent CTRs and the inverse kinematics problems, the developed infrastructure can, in real-time, leverage the dense workspaces for on-demand path-planning with on-line local inverse kinematics to guide the user along safe trajectories. Further, our research contributes to Implicit Active Constraints (IAC) for CTRs by proposing constraint generation based on the generated path-plans. Contrary to ACs, which are manually defined, IACs are algorithmically generated and tailored to the respective task based on limited user input [20]. As such, IACs are both more complex to define and to estimate intraoperatively, further increasing the benefit of a software architecture that enables real-time safe and effective robot guidance. Further expanding upon the work in the literature, the developed software requires minimal pre-computation, and is quick and reliable for incorporation in telemanipulation control schemes of CTRs. The computational benefits are supported by extensive quantitative evaluation. User-based evaluation of the guidance scheme is carried out through a study simulating a challenging clinical scenario from neurosurgery.

II. ROBOT KINEMATICS AND CONSTRAINTS

The sections of CTRs can be of variable curvature (VC) or fixed curvature (FC) [10]. A VC section consists of two tubes, which rotate individually but translate in tandem. Therefore, a VC section has 3 degrees-of-freedom (DoF). An FC section is a single tube and has 2 DoF. The \( i^{th} \) tube’s translation is denoted as \( \alpha^B \), and its base rotation as \( \alpha^B \). The joint-space \( q \) of a CTR is therefore described by the set of \( \{ \alpha^B \} \), where \( N_t \) is the number of tubes.

This work estimates the shape of the CTR based on the unloaded torsionally compliant kinematics model [4], [10]. Estimating the robot shape relies on computationally demanding iterative solving of a boundary value problem. Beginning at a torsion-free tip angle \( \alpha^T \), differential equations are iteratively solved via Euler approximation (discretisation step \( \epsilon_{arc} \)) to obtain the base angle \( \alpha^B \) for each tube. Subsequently, the shape of the robot is calculated using matrix exponentials and the curvature along the robot’s centre line.

A. Anatomical Constraints

CTRs are often deployed within delicate anatomy, making it important to avoid contact with the anatomy (denoted as \( \Gamma \)). Rasterized polygons, extracted from pre-operatively acquired images, provide the representation of \( \Gamma \) as a set of 3D points. Rasterisation occurs with a maximal lattice of \( \epsilon_{lat} \) and the points are stored in a \( k\)-d tree (3-d tree). This \( k\)-d tree allows for rapid distance queries between the robot shape and \( \Gamma \). Each point along the robot centreline is associated with the maximum tube diameter \( r \) at that robot location (the external surface of the robot). The \( k\)-d tree is queried to determine the nearest-neighbour \( \epsilon_{lat} \) of every robot point to \( \Gamma \). Given the tube radii, \( r^2 \), the discretisation step, \( \epsilon_{arc} \), and the mesh lattice, \( \epsilon_{lat} \), the distance to the anatomy, \( d_{ana} \), is:

\[
d_{ana} = \min_{i} \left( \epsilon_{lat} \epsilon_{nn}^2 - \frac{1}{2} \left\| \epsilon_{arc}, \epsilon_{lat} \right\|_2 - \epsilon_{lat} r^2 \right)
\]

A robot is colliding with the anatomy iff \( d_{ana} < 0 \).

B. Stability Constraints

Anatomical constraints are complemented by the mechanics-based risk of instability. Here, a quantitative measure of stability is used, first proposed in our work in [2]. In [10], it was demonstrated that instability manifests as an S-shaped curve relating the tube rotation at the base, \( \alpha^B \), to tube rotation at the tip, \( \alpha^T \). Observing Fig. 2, stable CTRs do not exhibit negative slopes on this S-curve [14]. This notion gives rise to the quantitative measure of stability, \( d_{sta} \):

\[
d_{sta} = \frac{\pi}{2} - \tan(\sigma^q)
\]

where \( \sigma^q \) is the minimum inverse slope along the entire S-curve. The angle difference between the most critical slope
of the S-curve $[\zeta = \arcsin(1, \sigma^2)]$ and the vertical axis ($\zeta = \frac{\pi}{2}$) represents the distance to an unstable configuration. The configuration is unstable if $d_{\text{sta}} < 0$.

The advantage of quantitative measurements for the anatomical constraints (distance to anatomy) and stability (distance to unstable configuration) is the ease of generation of differentiable cost-functions for inclusion in optimisation-based inverse kinematics solvers, and ii. thresholds for accepting/rejecting robot configurations under examination.

III. RAPID FORWARD AND INVERSE KINEMATICS

This section introduces and describes our high-performance software architecture to iteratively solve the boundary value problem defining the kinematics, while respecting stability and anatomical constraints. The developed software is based on the C++14 standard. It achieves a compromise between the multitude of criteria requiring consideration in high-performance robotics applications, e.g.:

1) memory allocation;
2) memory alignment and random memory access;
3) cache misses;
4) cache coherency;
5) branching, conditionals;
6) dynamic dispatch (virtual functions);
7) vector instructions.

Criteria 1-4 relate to memory access bottlenecks. Dynamically allocated memory is reserved on the heap, which is time consuming and should be avoided. Furthermore, all memory-access timings depend on where the data is stored, e.g. main memory, processor cache. Accessing the cache is order of magnitudes faster. Therefore, avoiding cache misses and increasing cache coherency helps increase execution speed.

Criteria 5-7 relate to bottlenecks regarding instructions. Branching means that the sequence of processor instructions depends on specific conditions that usually need to be evaluated at run-time. Delays are caused if the processor mispredicts the branch of instructions it needs to flush and therefore reloads the instruction pipeline. Dynamic dispatch (i.e. virtual functions) means that the instructions for a specific function call are determined at run-time instead of compile-time, which introduces an overhead for each function call. Finally, vector-instructions operate on an array of data simultaneously [single instruction multiple data (SIMD)] rather than on a single datum. Using vector-instructions in conjunction with expensive operations, such as trigonometric operations, can substantially speed-up algorithm execution.

Memory usage in CTRs mainly depends on the number of sample points along the robot centre line as well as the robot architecture (i.e. type of tubes and number of tubes). Fixing the number of sample points allows the preallocation of the maximally needed memory, and thus no allocation is necessary during the forward kinematic calculations. Further, the memory is constrained by memory-alignment requirements to enable vector instructions (e.g. 64-byte for the Xeon Phi™architecture). Respecting these requirements allows memory allocation as a single chunk, which simplifies and expedites memory copying and network transmission of robot objects. Further, optimal memory allocation and alignment helps to fully utilize cache lines and hence optimises overall processor cache utilization, which reduces cache misses and increases its coherency, reducing potential bottlenecks caused by memory access times.

Branching caused by conditional expressions is, when possible, reduced by substituting loops and conditionals with equivalent constructs that are resolved at compile-time. The next section describes our implementation for compile-time unrolled for-loops and if-else-statements, based on template meta-programming.

The advantage of compile-time resolved loops and conditionals is not only the avoidance of branching, but also the further optimisation (e.g. auto vectorization) of code instructions by the compiler, since most instructions can be fully determined at compile-time.

Dynamic dispatch in C++ is the usage of virtual function calls. The overhead of these calls can accumulate significantly in particular if the function’s complexity is minimal. We tackle this by avoiding some layers of polymorphism using variadic templates and precompiling many different concentric tube instances to avoid virtual function calls. Furthermore, vector instructions for costly functions (e.g.

```
1 #define CTR_ALIGNMENT 64
2 // Base class for the Concentric Tube Robot
3 template<class Real>
4 class alignas(CTR_ALIGNMENT) RobotBase
5 { }
6 public:
7 typedef RobotBase<Real> base_t;
8 virtual base_t* appS(const FC<Real>& _sec)
9 virtual VecJu calcKinematic(const VecJu _JVal,
10 const Real& _ArcLengthStep) noexcept = 0;
11 virtual Tr calcKinematic(const VecJe _JVal,
12 const Real& _ArcLengthStep) noexcept = 0;
13 }
```

Code 1: Abstract class representing the CTR architecture.
A variadic template class is marginal. The computational overhead resulting from dynamic dispatch class has only a limited interface of elevator functions so that components of a CTR, i.e. the index is a loop with dynamic loop count cannot be used; every

trigonometric functions or square-root) can be used since the structural description of the robot is known.

A. Robot architecture using template classes

Fully defining a robot description at compile time reduces adaptability for the software user, and a compromise between compile- and run-time definitions is required.

Using template classes as described below provides a fully defined robot kinematic architecture at compile time. Although compilation might require some extra time, the gain of shifting run-time computations to compile-time is significant. Furthermore, code inlining and compiled code optimisation is improved, since the sequence of instructions is transparent at the time of compilation.

We create an abstract base class (see Code 1) from which a variadic template class derives (see Code 3). The abstract class has only a limited interface of elevator functions so that the computational overhead resulting from dynamic dispatch is marginal. The Robot-class encapsulates the generating components of a CTR, i.e. its sections and section types (FC, VC) in a tuple (Code 3, Line 29). Tuples provide an optimal memory footprint while allowing to store the sections without abstraction.

Many different classes derived from the templated Robot class will get compiled, covering several expected robot architectures\(^1\). For example if the maximum number of sections is set to 2, 6 different Robot classes will be automatically generated, corresponding to the 6 possible robot architectures maximally having 2 sections: \{[FC],[VC],[FC,FC],[VC,FC],[VC,VC]\}. The number of generated classes is \(n_{\text{class}} = 2^{(n_{\text{max,sec}}+1)} - 2\). For \(n_{\text{max,sec}} = 8\) the compiler generates 510 classes in less than a minute.

\(^1\)Compilation of multiple Robot classes is only necessary if the used robot architecture is not known at compile time. In case the software is tailored to a specific robot architecture abstraction or compilation of a set of differently templated classes is not necessary.

```cpp
1 // Helper struct to recursively append a section to the
2 template-class Real, class ...Args> /\'
3 #define CTR_ALIGNMENT 8
4 template-class Real, class ...Args> class Robot;
5 namespace details {
6 template<size_t secs_s> 7 struct appH {
8 { 9 typedef RobotBase<Real> base_t; 10 template-class Real, class SecT, class ...Args>
11 static base_t* appS(const RobotBase<Real>&... _rob_prev,
12 const SecT&... _sec) noexcept
13 { return new Robot<Real,...,SecT>(._rob_prev,...,sec); }
14 }; 15 template typename // Stop criteria for recursive appending
16 struct appH<CTR_MAXSECT-1> { 17 { 18 typedef RobotBase<Real> base_t; 19 template-class Real, class SecT, class ...Args>
20 static base_t* appS(const RobotBase<Real>&... _rob_prev,
21 const SecT&... _sec) noexcept
22 { return nullptr; }
23}; 24 }
25 }

Code 2: Helper struct to append FC/VC section to the CTR.
```

```cpp
1 template-class Real, class ...Args>
2 class alignas(CTR_ALIGNMENT) Robot : public RobotBase<Real> {
3 { 4 // Base class type 5 typedef RobotBase<Real> base_t;
6 // Tuple type holding the different sections.
7 typedef std::tuple<Args...&, > secs_t; 8 // Tuple size, number of sections
9 static constexpr size_t secs_s = std::tuple.size<secs_t>::value;
10 public:
11 // Constructor to append a section.
12 Robot(R&& _rob_prev, SecT&& _sec)
13 { 14 // arg: robot with one section less.
15 II arg: section to append. /\'
16 template-class Real, class SecT>
17 explicit Robot(R&& _rob_prev, SecT&& _sec) noexcept
18 { m.Secs = std::tuple.cat(_rob_prev.m.Secs,
19 std::forward_as_tuple(_sec));}
20 virtual base_t* appS(const FC<Real>& ...sec) const noexcept
21 { return details::appH<secs_s>::template
22 appS<Real,FC<Real>,Args...&>(_rob_prev,...,sec); }
23 appS<Real,VC<Real>,Args...&>(_rob_prev,...,sec);
24 virtual Tr calcKinematic(const VecJ, _JVal)
25 { return nullptr; }
26 // Different sections stored in tuple
27 alignas(CTR_ALIGNMENT) secs_t.m.Secs;
28 }

Code 3: CTR variadic template class. The template arguments determine the section types, order, and number.
```

The recursive generation of new classes stops when the upper limit of total section number is reached. Using partial-specialization for the template arguments (see Code 2, Line 16-24) the recursive instantiation process is halted. The number of sections and tubes for each generated class is known at compile time. Iterating over the tuple storing the section is also possible at compile time.

B. Unrolling loops at compile time

Iterating over tuples of sections requires the availability, at compile time, of the indices of the elements to be accessed, i.e. the index is a template-parameter. Therefore, a standard for-loop with dynamic loop count cannot be used; every

```cpp
1 template <int Iter, int Last, int Inc = 1>
2 struct static for 
3 { 4 template <typename Fn>
5 inline void operator() (Fn&& fn) const noexcept
6 { // Call function
7 fn[std::integral_constant<int, Iter>() ];
8 // Recursive call
9 static for<Iter+Inc>
10 // Last iteration if positive increments
11 [Inc<0>](Iter+Inc, (Iter+Last-1)/Inc))>
12 // Last iteration if negative increments
13 [Inc<0>](Iter+Inc, (Iter+Last-1)/-Inc)),
14 Inc<>()(std::forward<Fn>(fn));
15 } 16 }
17 }
18 template <int Last, int Inc>
19 struct static for<Last, Last, Inc>
20 { // Termination criteria
21 template <typename Fn>
22 inline void operator() (Fn&& fn) const noexcept{ }
23 }

Code 4: for-loop resolved at compile time, using template-ing and partial specialization.
Code 5: Usage of static_for using generic lambdas to provide the iteration index as template-argument.

A standard approach for this is to use recursive template calls, which act with each level of recursion on the next element in the tuple. This is addressed with a developed dedicated convenience struct: static_for (see Code 4). Passing a function-pointer to static_for results in a call to the pointed function with the current iteration index, and a recursive call to static_for with an incremented (Inc > 0) or decremented (Inc < 0) Iter-value. This approach unrolls the for-loop at compile-time, resulting in a speed-up of computations at runtime as the instructions can be vectorized. The lengthy calculations for the Last-template parameter are necessary to ensure that the stop criterion is met (i.e. if \((\text{Last-First}) \mod \text{Inc} \neq 0\)). To use the function parameter as a template parameter, we package the iteration index into an integral constant-class (see Code 4, Line 7).

Using generic lambdas (see Code 5, Line 3) we can retrieve the loop index in the function definition, from the type [decltype(ie)] of the function parameter as constexpr, since we encoded it in the parameter type, instead of the parameter value. Code 5 shows the usage and the output using the static_for construct. The output, listed below the code segment, shows that each element type is different.

### C. Resolving conditionals at compile-time

Runtime performance is commonly decreased by the abundance of machine-level jumps created by if/else statements. To improve performance by limiting the number of such jumps at compile time, an approach similar to the static_for-loop is pursued. When different code sequences need to be executed depending on the CTR section type, a dedicated construct, static_if_else, is used. Two function-parameters are passed to the static_if_else struct (Code 6). The first represents the if-instructions, while the second represents the else-instructions. Depending on the template-parameter of the static_if_else, the correct branch is already chosen at compile-time.

Using static_for and static_if_else to compile CTR classes for all expected combinations of sections, every single class can be individually optimized by the compiler, even if from a developer’s perspective only a single class needs to be written. The majority of machine-level code optimisations is achieved by the compiler, while the complexity of the implementation is hidden from the library user.

The next section leverages the speed-up achieved by the described implementation, detailed by benchmarks in Sec. VI-A, towards on-line path-planning and inverse kinematics for assisted telemanipulation.

### IV. GUIDANCE VIA IMPPLICIT ACTIVE CONSTRAINTS

Based on our high-performance software architecture, we generate IACs [20]. In our work, guiding paths (i.e. path plans) are generated based on pre-computed road-maps and navigation goal position as defined by the operating clinician. Due to the computational efficiency of the software framework pre-computation occurs in matter of minutes, rather than hours. In the pre-computation step, random configurations of the robot are generated using a parallel computing approach. A schematic representation of the framework is depicted in Fig. 3. Our path-planner’s efficiency heavily relies on the developed multi-node framework. A central computer (server) controls computing clients that generate random robot-configuration samples and solve the inverse kinematics using different optimization techniques (see Sec V).


#### A. Probabilistic Roadmap

The constraint-generation framework is based on an undirected graph \( G \), with vertices \( v_i \in G \). The vertices of
the effect of the scaling factor are reported in Sec. VI-B.

Extended robot-configurations in most scenarios will be task-space, the joint-space sampling has to be non-linear. The extent of sampling elongated versus retracted robots is governed by $\gamma_{\varphi}$, as depicted in Fig. 4.

Each client calculates the forward kinematics for the given joint value, determines $d_{\text{sta}}$ (2) and $d_{\text{ana}}$ (1), and sends the configuration to the server iff

$$d_{\text{col}}(q_r) \leq d_{\text{col}}^{\text{thres}} \land d_{\text{sta}}(q_r) \leq d_{\text{sta}}^{\text{thres}}.$$  

Results on the time required to compute the road-maps and the effect of the scaling factor are reported in Sec. VI-B.

When the server has received a minimum number of configurations, a local planner calculates edges $e_{i,j}$ and the corresponding cost for transitioning from one robot configuration represented by the vertex $v_i$ to another vertex $v_j$. Edge generation is performed in two stages: first the cost for a potential edge $e_{i,j}$ has to be below a certain threshold, and second, only the edges with the $N_{cl}^\text{max}$ smallest costs are introduced in the graph. The edge $\omega(v_j, v_k)$ cost depends on

$$f_{ee}(v_j, v_k) = \|x_{ee}(v_j) - x_{ee}(v_k)\|$$

$$g_{jv}(v_j, v_k) = \|q(v_j) - q(v_k)\|$$

$$h_{cl}(v_j, v_k) = \sqrt{\sum_{i=1}^{N_{cl}} \|x_{cl}\{i, v_j\} - x_{cl}\{i, v_k\}\|^2}$$

where $x_{ee}$ provides the end-effector position of a vertex, $q$ returns the joint-value vector of vertex, $abs$ provides the element-wise absolute values of a vector, $N_{cl}$ is the larger of the centre line point counts of $v_j$, $v_k$, and $x_{cl}\{i, v_j\}$ provides the position of $i$-th centre line point of vertex $v_j$ if $i$ is smaller than the number of centre line points, and otherwise provides the position of the final centre line point.

The edge cost is calculated as:

$$\omega(v_j, v_k) = \omega(e_{j,k})$$

$$= f_{ee}(v_j, v_k) + \omega_{jv}^T g_{jv}(v_j, v_k) + \omega_{cl} h_{cl}(v_j, v_k)$$

where $\omega_{jv}^T$ is the vector of weights for the joint value differences, and $\omega_{cl}$ is the weight for the root-mean-square error of the centre-line differences. The graph generation process is detailed in [2].

C. Guidance Generation - Query of Roadmap

The guidance constraint is calculated from the implicit constraint defined by the desired user end-effector position and the roadmaps, utilizing a shortest-path search. The source vertex is the vertex of graph $G$ that is closest to the current configuration based on the cost-function described in (5.1)-(5.3). The goal vertex is determined by the user-defined target position. It is selected from a set of vertices that are ranked based on the distance between graph $G$ vertices and the target position. Since the computational cost of finding the shortest

Fig. 5. Implicit Active Constraints: Guidance paths to navigation goals.
path from the source vertex to all vertices in the goal set can be very high, we chose a subset of the potential vertices. During parallel computation, calculating paths to very similar goal configurations is avoided by maximising the cost, i.e. the robot configuration discrepancy, between the potential goal vertices. This increases the chances of finding a low-cost path from the source vertex. In a last step, multiple shortest path searches to the set of goal vertices are performed using the $A^*$ algorithm and the resulting lowest-cost path ($P_\ast$) is chosen for user guidance. An example of shortest paths to various targets is depicted in Fig. 5.

V. INVERSE KINEMATICS

The shortest-path search between the current robot configuration and the desired navigation target provides visual and haptic guidance to the user. Nevertheless, we ensure that the operator is in full control of the robot. Towards this end, we use the on-line inverse kinematics algorithm presented in [1] to allow user-intended deviation from the given paths. Multiple optimisation algorithms, namely:

- Controlled Random Search with Local Mutation;
- Multi-Level Single-Linkage with Low-Discrepancy;
- Bound Optimisation by Quadratic Approximation;
- Principal Axis;
- Software for unconstrained minimisation w/o derivative;
- Nelder-Mead simplex algorithm;
- Damped-least squares Jacobian;
- Controlled Random Search with Local Mutation;
- Multi-Level Single-Linkage with Low-Discrepancy;
- Principal Axis;
- Software for unconstrained minimisation w/o derivative;
- Nelder-Mead simplex algorithm;
- Damped-least squares Jacobian;

run in parallel on disparate computing devices to minimise the cost function:

$$C(q, T, \Gamma, p^D, z^D) = \gamma_{\text{pos}} |p^D - p^O|^2 + \gamma_{\text{ori}} \arcsin \left( \frac{|z^D - z^O|^2}{2} \right) + \gamma_{\text{ana}} C_{\text{ana}}(q, T, \Gamma) + \gamma_{\text{sta}} C_{\text{sta}}(q, T)$$  \hspace{1cm} (7)

This cost function depends on the current robot end-effector pose ($p^D, z^O$), anatomy ($\Gamma$), robot kinematics ($T$), and desired end-effector pose ($p^D, z^D$). Further, there is dependence on cost functions that encode risk of collision with the anatomy ($C_{\text{ana}}$) and configuration instability ($C_{\text{sta}}$).

The inverse kinematics solver is initialized based on (5.1)-(5.3) with the joint values of the vertex $v_i \in P_\ast$ closest to the current robot configuration. Therefore, finding a locally best configuration for the current desired position (tele-manipulation position) is based on the globally optimized configuration to reach the goal robot configuration.

Haptic guidance is based on an elasto-plastic friction model, which redirects kinetic energy from user input towards the guidance path [16]. The parameters governing the extent and feel of haptic guidance can be found in Table I and follow the formulas introduced in [1].

### TABLE I

<table>
<thead>
<tr>
<th>Parameter (GVF) Value</th>
<th>Parameter (GVF) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_C$ N</td>
<td>$\sigma_0$ N/mm</td>
</tr>
<tr>
<td>$\Theta_{\text{approx}}$</td>
<td>$\sigma_1$ N/s/mm</td>
</tr>
<tr>
<td>$z_{\text{css}}$ mm</td>
<td>$\sigma_2$ N/s/mm</td>
</tr>
</tbody>
</table>

### VI. EXPERIMENTAL RESULTS

A. Computational Benchmark

This section compares three kinematic implementations written in C++ to demonstrate the computational efficiency gained by the novelties of the work described herein.

The first implementation is «completely dynamic» and was used in [14]. Each robot consists of a set of sections, and each section consists of a number of tubes. Limited compile-time optimisation is possible with this approach, and, hence, the computational efficiency is low.

The second implementation, «simple template», encodes the robot as a template class with two parameters: i. number of sections, and ii. number of tubes. Although, this approach allows to fully unroll loops and to resolve conditional expressions at compile-time, the section-class had to be implemented using virtual-functions, which resulted in important run-time overheads.

The final approach, «variadic template», is the contribution of this paper and follows the new design patterns.

All benchmarks were performed on an Intel® Core™ i7-3770 CPU. Table II show profiler results from the Cachegrind-software package, when compiled with the Intel® C++ Compiler (ICC) 2016.3. It shows improvements of more than an order of magnitude between the template-code and the «completely dynamic» code. The improvements between the «variadic template» code and the «simple template» code are also significant, ranging from 13.7% (instruction references) to 42.4% (last-level cache references). Therefore, it is expected that computation time will significantly improve for the «variadic template» implementation.

The computation times for the three implementations, listed in Table III, validate this expectation. The computational-time reduction between the «completely dynamic» implementation and the template-code is more than an order of magnitude. Depending on the compiler choice, «variadic template» code ran up to 18.5 times faster. In comparison with the «simple template» code, an improvement of

### TABLE II

<table>
<thead>
<tr>
<th>Profiler metric</th>
<th>variadic template</th>
<th>simple template</th>
<th>completely dynamic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction refs.</td>
<td>112465</td>
<td>130365</td>
<td>192965</td>
</tr>
<tr>
<td>Data refs.</td>
<td>52665</td>
<td>67465</td>
<td>782465</td>
</tr>
<tr>
<td>Last-Level cache</td>
<td>0.3165</td>
<td>0.5565</td>
<td>2.7565</td>
</tr>
<tr>
<td>Branches</td>
<td>2565</td>
<td>4085</td>
<td>294965</td>
</tr>
<tr>
<td>Misprediction</td>
<td>0.2165</td>
<td>0.2565</td>
<td>5.1665</td>
</tr>
<tr>
<td>Misprediction rate</td>
<td>0.73%</td>
<td>0.63%</td>
<td>1.75%</td>
</tr>
</tbody>
</table>

### TABLE III

<table>
<thead>
<tr>
<th>Implementation</th>
<th>Time ICC 2016.3</th>
<th>Time Ratio*</th>
<th>Time ICC 6.1</th>
<th>Time Ratio*</th>
<th>Time Clang 3.8</th>
<th>Time Ratio*</th>
</tr>
</thead>
<tbody>
<tr>
<td>variadic template</td>
<td>1.75s</td>
<td>100%</td>
<td>2.94s</td>
<td>100%</td>
<td>2.80s</td>
<td>100%</td>
</tr>
<tr>
<td>simple template</td>
<td>1.88s</td>
<td>107%</td>
<td>3.25s</td>
<td>111%</td>
<td>2.93s</td>
<td>105%</td>
</tr>
<tr>
<td>completely dynamic</td>
<td>27.68s</td>
<td>1580%</td>
<td>32.28s</td>
<td>1100%</td>
<td>31.58s</td>
<td>1129%</td>
</tr>
</tbody>
</table>

*Ratio of the time compared to the respective fastest algorithm.
7% is measured using the ICC compiler. This evolutionary improvement justifies an increased initial implementation complexity because usability, and computational precision are identical and the source-code maintenance became simpler using the «variadic template». The compiler flags were tuned individually for all examined compilers to guarantee that each one performs at its best. The binaries produced by GCC 6.1 and Clang 3.8 had a similar runtime, whereas the ICC 2016.3 compiler produced a significantly faster code.

The relative improvements from ICC to GCC or Clang were less for the «completely dynamic» implementation.

B. Creation of Road-map

Durations for computing safe configuration samples for the road maps are listed in Table IV. The results show timings and acceptance rates for the respective translational scaling factor ($\gamma_x$). Obtaining a graph with 1,048,576 vertices, and 209,715 vertices per $\gamma_x \in [0.12, 0.25, 0.35, 0.70, 0.90]$ took 186.8 s with 149.4 s for the sampling and 37.4 s for defining the edges of the graph in the local planning step.

C. User Experiment

This section describes the benefits of the developed framework from the operator’s perspective, evaluating the value of interactive-rate inverse kinematics and implicit active constraints. The simulated clinical scenario is based on an intervention that involves cauterisation of the choroid plexus in hydrocephalic ventricles. In this procedure, an elongated CTR needs to access the base of the ventricles and cauterise them to limit the production of cerebrospinal fluid as a means to indirectly reduce ventricle pressure. The procedure is envisioned as an endoscopic to third ventriculostomy [7], and requires highly curved CTR that are prone both to instabilities and to collisions with the anatomy (see robot parameters in Table V).

In the experiment 4 novice female and 10 novice male participants, age 25-35, had to manoeuvre the CTR towards 13 cauterization points using a haptic device (Geomagic Touch). The task was performed in two different modes: «guidance», «free». The order of the modes was randomised for every participant. The experimental protocol recorded metrics on a seven-level Likert scale, such as execution time, frustration, ease of telemanipulation etc. Each participant could activate the IAC on-demand in the «guidance» mode to evaluate the effectiveness of the proposed cooperative guidance framework which provides: visual guidance, haptic guidance, and path-planning aided inverse kinematics, compared to the «free» mode, where no path guidance was provided but only collision avoidance guidance and the inverse kinematics was solely based on on-line optimization. Users could choose to follow closely or approximatively the guidance path to command the robot, allowing easy manipulation even when the shape of the path appeared complex.

The list of quantitative performance metrics used includes:

- Duration [min]: Duration of task completion;
- $P_{ee}$ TD [m]: Travel distance of CTR tip;
- $P_{sp}$ TD [m]: Travel distance of CTR desired position;
- $\Delta P_{ee\leftrightarrow sp}$ O.T. [m]: Integral of $\| P_{ee} - P_{sp} \|^2$;
- $\Delta P_{IK}$ [mm]: Accuracy of inverse kinematics;
- $\Delta T_{opt}$ [ms]: Mean time of IK updates.

The results of the experiment are listed in Table VI. They show statistical significance based on the Wilcoxon signed-rank test in metrics 1-4. This indicates an overall better task performance when using the proposed path-planning and guidance framework. Fig. 6 shows two plots of the distance between the operator’s desired position and the CTR’s tip position indicating that the path-planning improves the convergence of the inverse-kinematics; the reason for this is that without path-planning the local inverse kinematics solver cannot escape a local minimum, which originates from the stability constraints.

In a qualitative assessment, the participants were asked to compare the modes, on a scale from $-3$ to $+3$ ($-3$: negative evaluation, $+3$: positive evaluation). The results are listed in Table VII. The majority of the users stated to prefer the «guidance» mode over the «free» mode. The users felt more in-control with guidance enabled, giving them a safer feeling. Unsurprisingly, the haptic feedback arising from the attractive and repulsive elasto-plastic friction model was evaluated as useful by the users.

### Table IV

<table>
<thead>
<tr>
<th>Scaling ($\gamma_x$)</th>
<th>configurations 250,000</th>
<th>configurations 1,000,000</th>
<th>Parametrization: $\gamma_{arc,ik}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>107.98 s</td>
<td>550.83 s</td>
<td>1 mm</td>
</tr>
<tr>
<td>0.4</td>
<td>213.5 s</td>
<td>811.3 s</td>
<td>1 mm</td>
</tr>
<tr>
<td>0.7</td>
<td>163.9 s</td>
<td>704.8 s</td>
<td>2 mm</td>
</tr>
<tr>
<td>1.0</td>
<td>174.8 s</td>
<td>603.3 s</td>
<td>1 mm / 5 mm</td>
</tr>
</tbody>
</table>

### Table V

<table>
<thead>
<tr>
<th>Robot Section Type</th>
<th>Curvature [mm$^{-1}$]</th>
<th>Straight/Curved Length [mm]</th>
<th>Stiffness Ratio</th>
<th>Tube</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>variable curvature</td>
<td>0.0522</td>
<td>0.00/66.47</td>
<td>5.1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>fixed curvature</td>
<td>0.05263</td>
<td>66.47/66.70</td>
<td>1:1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>fixed curvature</td>
<td>0.0</td>
<td>113.17/12.65</td>
<td>0.05:1</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

### Table VI

<table>
<thead>
<tr>
<th>Duration [min]</th>
<th>Free Mode</th>
<th>Guidance Mode</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.9 [10.5,15.9]</td>
<td>8.6 [6.0,10.5]</td>
<td>0.9%</td>
<td></td>
</tr>
<tr>
<td>$P_{ee}$ TD [m]</td>
<td>2.0 [1.8, 3.5]</td>
<td>1.1 [1.0, 1.3]</td>
<td>1.6%</td>
</tr>
<tr>
<td>$P_{sp}$ TD [m]</td>
<td>3.2 [2.7, 4.7]</td>
<td>1.4 [0.9, 1.6]</td>
<td>0.00%</td>
</tr>
<tr>
<td>$\Delta P_{ee\leftrightarrow sp}$ O.T. [m]</td>
<td>7.4 [2.9, 8.7]</td>
<td>0.6 [0.5, 1.4]</td>
<td>1.6%</td>
</tr>
<tr>
<td>$\Delta P_{IK}$ [mm]</td>
<td>0.7 [0.4, 1.7]</td>
<td>0.7 [0.2, 0.8]</td>
<td>11.3%</td>
</tr>
<tr>
<td>$\Delta T_{opt}$ [ms]</td>
<td>2.0 [2.0, 2.0]</td>
<td>2.0 [2.0, 2.0]</td>
<td>20.6%</td>
</tr>
</tbody>
</table>

$^1$Value represents the median, in parenthesis 25th and 75th percentile.
The paper presents a novel software design and computing architecture to rapidly calculate dense path-plans to safely guide the operator during telemanipulation of a CTR. The emphasis of this paper is on the high-level code optimisation architecture to rapidly calculate dense path-plans to safely reduce it further by investigating approaches that warp the time pre-computation is required, and by the time required in software engineering, as well as via a multi-user approach is demonstrated by a variety of metrics established in the literature highlight the inverse kinematics of CTR as a highly resource-intensive problem, which our software architecture makes tractable and solvable in real-time. The performance and effectiveness of the presented approach is demonstrated by a variety of metrics established in software engineering, as well as via a multi-user experiment that focused on human factors and useability.

Our approach is still limited by the fact that a reduced-time pre-computation is required, and by the time required to calculate the path plan intraoperatively. Even though this amounts only to 3-5 min, and 1-10 s, respectively, we aim to reduce it further by investigating approaches that warp the workspace based on real-time collision avoidance to enable dynamic active constraints, providing operator guidance in cases where the anatomy undergoes motion and manipulation. Without frameworks such as the proposed one, advances in complex-robot telemanipulation will lag.

The proposed approach of completely templated classes can also be used for other variations of continuum robots where the kinematics variables and robot shape are coupled, leading to computationally intensive kinematics. Kinematic chains of standard serial-link robots could also be an area of application, with each tuple element representing a different type of joint (e.g. rotational, prismatic, universal), albeit with limited benefit due to their significantly simpler kinematics.

### VII. DISCUSSION AND CONCLUSIONS

The paper presents a novel software design and computing architecture to rapidly calculate dense path-plans to safely guide the operator during telemanipulation of a CTR. The emphasis of this paper is on the high-level code optimisation architecture to rapidly calculate dense path-plans to safely reduce it further by investigating approaches that warp the time pre-computation is required, and by the time required in software engineering, as well as via a multi-user approach is demonstrated by a variety of metrics established in the literature highlight the inverse kinematics of CTR as a highly resource-intensive problem, which our software architecture makes tractable and solvable in real-time. The performance and effectiveness of the presented approach is demonstrated by a variety of metrics established in software engineering, as well as via a multi-user experiment that focused on human factors and useability.

### VIII. ACKNOWLEDGEMENT

The authors gratefully acknowledge Prof. Pierre Dupont from Harvard Medical School for providing the anatomy used in this paper.

### REFERENCES


