Poly-Time Knot Theory and Quantum Algebra — Discovery Grant Proposal

Recently, Roland van der Veen and myself, following Lev Rozansky and Andrea Overbay [Ro1, Ro2, Ro3, Ov], presented a methodology [BV2] for the construction of poly-time computable knot polynomials and constructed [BV1] the first poly-time computable knot polynomial since the Alexander polynomial of 1928 [Al].

Why is it exciting, even before the details? Here and there mathematics has immense philosophical value or beauty to justify the effort. Yet everyday mathematics is mostly about, or should be about, "doing useful things". Deciding if A has property B, counting how many C's satisfy D, computing E. When A and B and C and D and E are small, we do the computations on the back of an envelope and publish them as "Example 3.14" in some paper. But these are merely the demos, and sooner or later we worry (or ought to worry) about bigger inputs. I'm more aware then most mathematicians (though perhaps less than many computer scientists), how much the complexity of obtaining the solution as a function of the size of the inputs matters. Hence I firmly believe that incomputable mathematics is intrinsically less valuable than computable mathematics (allowing some exceptions for philosophical value and/or beauty), and that within computable mathematics, what can be computed in linear time is generally more valuable than what can be computed in polynomial (poly-) time, which in itself is more valuable than what can be computed in exponential (exp-) time, which in itself is more valuable than what can be computed just in theory.

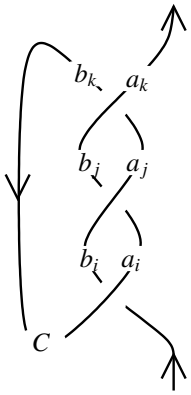
With the exception of the Alexander polynomial, which has been thoroughly mined[1], and with the exception of the first few finite-type invariants [BN1, BN2], which are rather weak, until recently [BN8] all known knot invariants were harder-than-poly-time to compute[2]. So clearly, what we have in [BV1, BV2] — a rather strong poly-time-computable knot polynomial[3] and a methodology for further such — is a priori exciting.

Furthermore, our invariants extend to tangles, and are well-behaved under the basic tangle-theoretic operations of "strand stitching" and "strand doubling", and hence they carry topological information: a bound on the genus of a knot [BN9], and what may be the best chance we have of showing that certain slice knots are not ribbon [BN11], hence resolving (negatively) the long-standing "slice=ribbon" conjecture [FM].

Finally, merely our suggestion (starting [BN7]) that some poly-time knot polynomials beyond the Alexander polynomial ought to exist is already generating both interest [Fi] and competition [Pr] (both authors do not cite us explicitly, but were present in our talks [BN12, BN10] and were clearly influenced).

Our methodology. Since already the 1980s, there is a standard "quantum algebra" methodology that associates a framed knot invariant to certain triples (U, R, C), where U is a unital algebra and $R \in U \otimes U$ and $C \in U$ are invertible (see e.g. [Oh]). Let us briefly recall the standard methodology here.

Draw a given (framed, oriented) knot K as a long knot in the plane so that at each crossing the two crossing strands are oriented upward, and so that the orientation at two ends of K is up.



$$z(K) = \sum_{i,j,k} b_i a_j b_k C a_i b_j a_k$$

Put a copy of $R = \sum a_i \otimes b_i$ on every positive crossing of K with the "a" side on the over-strand and the "b" side on the under-strand, labeling these a's and b's with distinct indices $i, j, k, \ldots$ (similarly put copies of $R^{-1} = \sum a'_i \otimes b'_i$ on the negative crossings; these are absent in our example). Put a copy of $C^{\pm 1}$ on every cuap where the tangent to the knot is pointing to the right (meaning, a C on every such cup and a C^{-1} on every such cap).

Form an expression $z(K)$ in U by multiplying all the a, b, C letters as they are seen when traveling along K and then summing over all the indices, as shown.

If R and C satisfy some conditions dictated by the standard Reidemeister moves of knot theory, especially the Yang-Baxter Equation (YBE), the resulting $z(K)$ is a knot invariant.

Abstractly, $z(K)$ is obtained by tensoring together several copies of $R^{\pm 1} \in U^{\otimes 2}$ and $C^{\pm 1} \in U$ to get an in-

[1] Though newer and better still arises. For example, the techniques of [BNS] lead to the fastest known algorithm for the computation of the Alexander polynomial.

[2] Though divide-and-conquer methods reduce the computation time for the Jones and HOMFLY-PT polynomials [Jo, HOMFLY, PT] and possibly even for Khovanov homology [BN3] to around $C e^{c\sqrt{n}}$ where n is the crossing number, and so these invariants can be computed for surprisingly large knots.

[3] How strong? As detailed in [BN9], stronger than HOMFLY-PT and Khovanov taken together, at least for knots up to with 12 crossings.

termediate result $z_0 \in U^{\otimes S}$, where S is a finite set with two elements for each crossing of K and one element for each right-pointing cuap. We then multiply the different tensor factors in z_0 in an order dictated by K to get an output in a single copy of U.

The best algebras with which to apply the standard methodology, at least as of 2017, are certain completions $\mathcal{U}(\mathfrak{g})$ of the universal enveloping algebras of semi-simple Lie algebras $\mathfrak{g}$ (or their quantizations). But these algebras are infinite dimensional and the sum in $z(K)$ is infinite and not immediately computable.

The dogma solution is to pick a finite dimensional representation of $\mathfrak{g}$ and use it to represent all the elements appearing in $z(K)$, effectively replacing the algebra by the algebra of endomorphisms of some finite dimensional vector space. This turns the sum finite; yet if the knot K has n crossings, our sum becomes a sum over n indices $i_1, \dots, i_n$. Thus there are exponentially-many summands to consider and it takes an exponential amount of time to compute $z(K)$.[2,4]

Alternatively, one may extract finite-type [BN1] information out of $z(K)$ by reducing modulo appropriate filtrations of U and its tensor powers. As already mentioned, the results are computable but weak.

Our approach to the computation of $z(K)$ is different. Instead of working directly in $U^{\otimes S}$, we work in relatively small[5] spaces $\mathcal{F}(S)$ of "closed-form formulas for elements of $U^{\otimes S}$". For this to work, we need to ensure that the fundamentals R and C would be described by "closed-form formulas", and that the most basic operations necessary for the computation of z, namely multiplication of factors in $U^{\otimes S}$, would be implemented "in closed form" in $\mathcal{F}(S)$.

In practice, the kind of terms that appear within formulas for R and C are exponentials of the form $e^{\xi x}$, where x is a generator of U and ξ is a formal scalar variable, their iterated derivatives $(\partial_\xi)^k e^{\xi x} = x^k e^{\xi x}$, and exponentials of quadratics like $e^{\lambda xy}$ or $e^{\lambda x \otimes y}$, with scalar λ and $x, y \in U$. We then need to multiply several such exponentials and differentiated exponentials, and we need to learn how to bring such products into some pre-chosen "canonical order". In the standard $U \sim \mathcal{U}(\mathfrak{g})$ case, where $\mathfrak{g}$ is semi-simple, this is complicated. Yet if $\mathfrak{g}$ is solvable, this is often easy. Wouldn't it be nice if it was possible to approximate semi-simple Lie algebras with solvable ones?

In our work we exploit the little-known fact that this is (nearly) possible. Precisely, given a semisimple $\mathfrak{g}$, there exists a Lie algebra $\mathfrak{g}^\epsilon$ defined over the ring $\mathbb{Q}[\epsilon]$ of polynomials in a formal variable ϵ (in other words, $\mathfrak{g}^\epsilon$ is a "one-parameter family of Lie algebras"), so that

1. If ϵ is fixed to be some constant not equal to zero, then $\mathfrak{g}^\epsilon$ is isomorphic to $\mathfrak{g}^+ := \mathfrak{g} \oplus \mathfrak{h}$, which is the original $\mathfrak{g}$ with an extra copy of its own (Abelian) Cartan subalgebra $\mathfrak{h}$ added.

2. At $\epsilon = 0$, $\mathfrak{g}^0$ is solvable. Furthermore, $\mathfrak{g}^\epsilon$ is solvable in a formal neighborhood of $\epsilon = 0$: for any natural number $k \geq 0$ the reduction $\mathfrak{g}^{\leq k}$ of $\mathfrak{g}^\epsilon$ to the ring $\mathbb{Q}[\epsilon]/(\epsilon^{k+1} = 0)$ is solvable as a Lie algebra over $\mathbb{Q}$ (whose dimension is $(k+1)\dim\mathfrak{g}$).

As k gets larger, the solvable $\mathfrak{g}^{\leq k}$ is closer and closer to $\mathfrak{g}^\epsilon$, as the reduction modulo $\epsilon^{k+1} = 0$ means less and less, and so at least informally, $\mathfrak{g}^{\leq k} \xrightarrow[k\to\infty]{} \mathfrak{g}^+ \sim \mathfrak{g}$.

We sketch why $\mathfrak{g}^\epsilon$ exists.

Let $\mathfrak{g}$ be a semisimple Lie algebra and let $\mathfrak{b}^+$ and $\mathfrak{b}^-$ be its upper and lower Borel subalgebras, respectively. Then $(\mathfrak{b}^+)^*$ is $\mathfrak{b}^-$, and as the latter has a Lie bracket, it follows that $\mathfrak{b}^+$ has a co-bracket δ. In fact, $\mathfrak{b}^+$ along with its bracket $[\cdot,\cdot]$ and co-bracket δ is a "Lie bialgebra", and one may recover $\mathfrak{g}^+ = \mathfrak{g} \oplus \mathfrak{h} = \mathfrak{b}^- \oplus \mathfrak{b}^+$ as the "Drinfel'd double" $\mathcal{D}(\mathfrak{b}^+, [\cdot,\cdot], \delta)$ of $\mathfrak{b}^+$ (see e.g. [ES]). The axioms of a Lie bialgebra are homogeneous in δ, meaning that $(\mathfrak{b}^+, [\cdot,\cdot], \epsilon\delta)$ is again a Lie bialgebra for any scalar ϵ, and one may set $\mathfrak{g}^\epsilon := \mathcal{D}(\mathfrak{b}^+, [\cdot,\cdot], \epsilon\delta)$. The required properties are easy to check. Perhaps the most interesting is the solvability of $\mathfrak{g}^0$: indeed $\mathfrak{g}^0 = I\mathfrak{b}^+ := (\mathfrak{b}^+)^* \rtimes \mathfrak{b}^+$ with $(\mathfrak{b}^+)^*$ regarded as an Abelian Lie algebra and $\mathfrak{b}^+$ acting on $(\mathfrak{b}^+)^*$ using the co-adjoint action, and then the solvability of $I\mathfrak{b}^+$ easily follows from the solvability of $\mathfrak{b}^+$. We studied the knot-theoretic significance of $\mathfrak{b}^* \rtimes \mathfrak{b}$ for a general Lie algebra $\mathfrak{b}$ extensively in the context of "w-knots" in [BND1, BND2, BND3, BN6, BN5] (supported by our previous NSERC discovery grant), and these studies along with the observations in this paragraph were in some sense the starting points for our current study.

A model result. Let us state the $\mathfrak{g} = sl_2$ version of our main result; one of our future directions will be to extend beyond that case. Let k be a natural number.

One may check that the Lie algebra $sl_2^{\geq k}$ is generated by generators $\{t, y, a, x\}$ such that t is central and $[a,x] = x$, $[a,y] = -y$, and $[x,y] = 2\epsilon a - t$; here ϵ is

[4] Note that almost any time the phrases "braided monoidal category" or "TQFT" are used within low dimensional topology, some tensor powers of some vector spaces need to be considered at some point, and dimensions grow exponentially. Thus our criticism applies in these cases too [BN12].

[5] Ranks grow polynomially in $|S|$.

a scalar for which $\epsilon^{k+1}=0$. Let $\mathcal{U}$ be the universal enveloping algebra of $sl_2^{\le k}$ (completed, details elsewhere), and likewise let $\mathcal{S}$ be the (similarly completed) symmetric algebra of $sl_2^{\le k}$. We write "v_i" for "v placed in an i'th tensor factor", so e.g., $y_1x_2 = y\otimes x$ is an element of either $\mathcal{U}^{\otimes 2}$ or $\mathcal{S}^{\otimes 2}$, according to context. We let $\mathbb{O}\colon \mathcal{S}\to\mathcal{U}$ be the "normal ordering map" which plants the non-commuting generators $\{y,a,x\}$ in the "y then a then x" order. For example, for a scalar λ, $\mathbb{O}(e^{\lambda xy}) = \sum \frac{\lambda^j y^j x^j}{j!}$ (and not $\sum \frac{\lambda^j (yx)^j}{j!}$).

It turns out that there are suitable R and C elements for $\mathcal{U}$, constructed using the standard quantum algebra methodology, and hence there is a corresponding knot invariant z, which easily extends to some general class of tangles which we do not specify here other than to say that it includes all knots.

Model Theorem. *If K is an n-crossing S-component tangle (where S is a finite set) then $z(K)\in\mathcal{U}^{\otimes S}$ and*

$$z(K)=\mathbb{O}\left(\omega\exp\left(\sum_{i,j\in S}\lambda_{ij}t_ia_j+\sum_{i,j\in S}q_{ij}y_ix_j+\sum_{d=1}^{k}P_d\epsilon^d\right)\right),$$

where $\lambda_{ij}\in\mathbb{Z}$, where ω and q_{ij} are rational functions in $T_i := e^{t_i}$ with numerators and denominators of degrees bounded by n, and where the P_d's are "perturbations" which are polynomials in variables $\{y_i,a_i,x_i\}$ of degree at most $4d$, with coefficients rational functions in the T_i's with numerators and denominators of degrees bounded by n.[6]

The formula for $z(K)$ appears complicated, but in some technical sense, it is in fact simple. Indeed if all the T_i are identified (setting $T_i=T_j$ for all $i,j\in S$), then the space $\mathcal{F}(S)$ of all formulas as in the theorem is of "poly-size" — such a formula is determined by a finite number of integer coefficients and the number of coefficients required is a polynomial in n and in $|S|$.

The fact that $\mathcal{F}(S)$ is poly-size suggests that the operations one needs to perform on $\mathcal{F}(S)$ to compute $z(K)$ (mostly, multiplication of different tensor factors) would take poly-time. We show that this is indeed the case, and hence $z(K)$ is poly-time computable.

What goes in the proof? Some "classical algebra" PBW-reordering techniques to carry out the required operations on $\mathcal{F}(S)$, some "quantum algebra" techniques to find R and C in a certain quantization of $\mathcal{U}$, and a little bit of extra work to pull R and C from the quantized world to the classical one, in which our model theorem is written.

In the case where $k=1$, this isn't just "theory" — the programs are written and are quite short, and the results are tabulated and are quite strong; see e.g. [BN10, BV1]. For $k>1$ we are very near a complete implementation. It is not much more complicated, and the results should be more powerful. The generalization to $\mathfrak{g}$ other than sl_2 is in sight, yet will require more work.

Frequently Asked Questions.

- *What is the relationship between this proposal and the work of Rozansky and Overbay?* Our invariants are the Rozansky-Overbay [Ro1, Ro2, Ro3, Ov] invariants, and in many ways our work is a continuation of theirs, though we believe our techniques are cleaner and more easily susceptible to generalization. Rozansky and Overbay did not note that their invariants are computable in polynomial time, and did not explain how to generalize them to the case of tangles. The latter generalization allows for divide-and-conquer computations which lead to a significant speedup, and is crucial if one attmpts to relate invariants to topological properties such as the genus and the ribbon property (e.g. [BN11]).

- *What is the relationship between these invariants and the coloured Jones polynomial?* The totality of all $sl_2^{\le k}$ invariants is most likely equivalent to the coloured Jones polynomial; incrementing k by one should correspond to the consideration of one further diagonal in the Melvin-Morton-Rozansky expansion of the coloured Jones polynomial [MM, BNG]. In some sense, as indicated already in the "summary" section of this proposal, we merely find a computable "corner" of a known and very-difficult-to-compute theory. We believe computability makes a huge difference! Note also that there isn't a good Melvin-Morton-Rozansky expansion for tangles, and tangles are crucial from our perspective.

- *What is the relationship between this proposal and the "loop expansion" of the Kontsevich integral?* The "loop expansion" [GR] of the Kontsevich integral is an all-Lie-algebra universal version of the Melvin-Morton-Rozansky expansion, and it should relate to the invariants we discuss in accordance with that — for any semi-simple $\mathfrak{g}$, the $\mathfrak{g}^{\le k}$ should be "k-loop invariants". Yet again, the "loop expansion" is nearly impossible to compute and its

[6] The actual degree bounds we have are stronger than indicated here, though a bit harder to state. Stronger degree bounds imply faster computations.

generalization to tangles depends on a choice of a Drinfel'd associator, which is another hard-to-compute object.

Our proposed research. Much remains to be done, and I plan to do it over the grant period:

- Our present implementation of the algorithm for the $sl_2^{\leq 1}$ invariant is pathetic; it has the right asymptotic behavior, but the constants are all wrong, by factors of thousands. This should be improved.
- The implementation work in the case of $k > 1$ has to be completed.
- An excellent exposition for the case $\mathfrak{g} = sl_2$ and for general k should be written.
- Everything should be generalized and implemented for Lie algebras beyond sl_2. If $\mathfrak{g}$ is of rank r, meaning its Cartan subalgebra $\mathfrak{h}$ is r-dimensional, then our invariants should become a (computable) polynomials in r variables!
- Our tabulations so far (e.g. [BN10]) show that the $k = 1$ invariant for sl_2, $\rho_1(K)$, yields a bound on the genus $g(K)$ of K: it seems that always $\deg \rho_1(K) \leq 2g(K) - 1$, and that that bound is independent of the bound obtained from the Alexander polynomial. I think I know why this is so, and why there should be a "Seifert surface formula" for $\rho_1(K)$, but this has to be rigorously confirmed.
- As indicated in [BN11], there is a potential for a relationship between these invariants, in particular ρ_1, and ribbon knots and the ribbon=slice conjecture. This should be pursued.
- There is some tension within our work between classical universal enveloping algebras and quantized ones: operations of $\mathcal{F}(S)$ are easier to describe in classical language, yet the crucial elements R and C are easier to describe in the quantum language. Our current solution is to use an isomorphism between the two languages (an analog of the non-canonical algebra-structure-only isomorphism $\mathcal{U}(\mathfrak{g})[\![\hbar]\!] \simeq \mathcal{U}_\hbar(\mathfrak{g})$) to push/pull structures from one side to the other. We expect that a better understanding of this tension will eventually arise, and with it a better understanding of quantum groups as they appear in knot theory.[7]
- The relationship between the story presented here and the "loop expansion" of the Kontsevich integral (e.g. [GR]) should be studied.
- More generally, there may be more to say about poly-time computations in knot theory (e.g. [Pr, Fi]). I intend to contribute in these directions as well.
- While these topics are largely untouched within this proposal, I intend to continue working as time will allow on the topics of my previous NSERC research proposal, "Knot Theory, Algebra, and Higher Algebra", [BN4, BNS, BN5, BND1, BND2, BND3, BN6].

[7] Perhaps a footnote isn't the right place to raise a major issue, yet I believe we don't genuinely understand the relationship between quantum groups and knot theory: if we know quantum groups we know how to make knot invariants, but the relationship should also go the other way. One has to be able to start with "we want knot invariants" and be lead to the specific formulas appearing in the quantizations of semi-simple Lie algebras. The narrative for the latter direction, as it stands now, is far from complete. It is hard to expect that our work in itself will change this situation. Yet we must hover around these issues if we ever want to fully understand them, and this we do.