✨ Member-only story

YEAR 2021 IN REVIEW & PREDICTIONS FOR 2022

What does 2022 hold for Geometric & Graph ML?

Last year, I sought the opinion of leading researchers of Graph ML to make predictions about the future development in the field. This year, we teamed up with Petar Veličković and interviewed a cohort of distinguished and prolific experts in an attempt to summarise the highlights of the past year and predict what is in store for 2022.

Michael Bronstein · Following
Published in Towards Data Science
44 min read · Jan 24, 2022

Listen Share ••• More

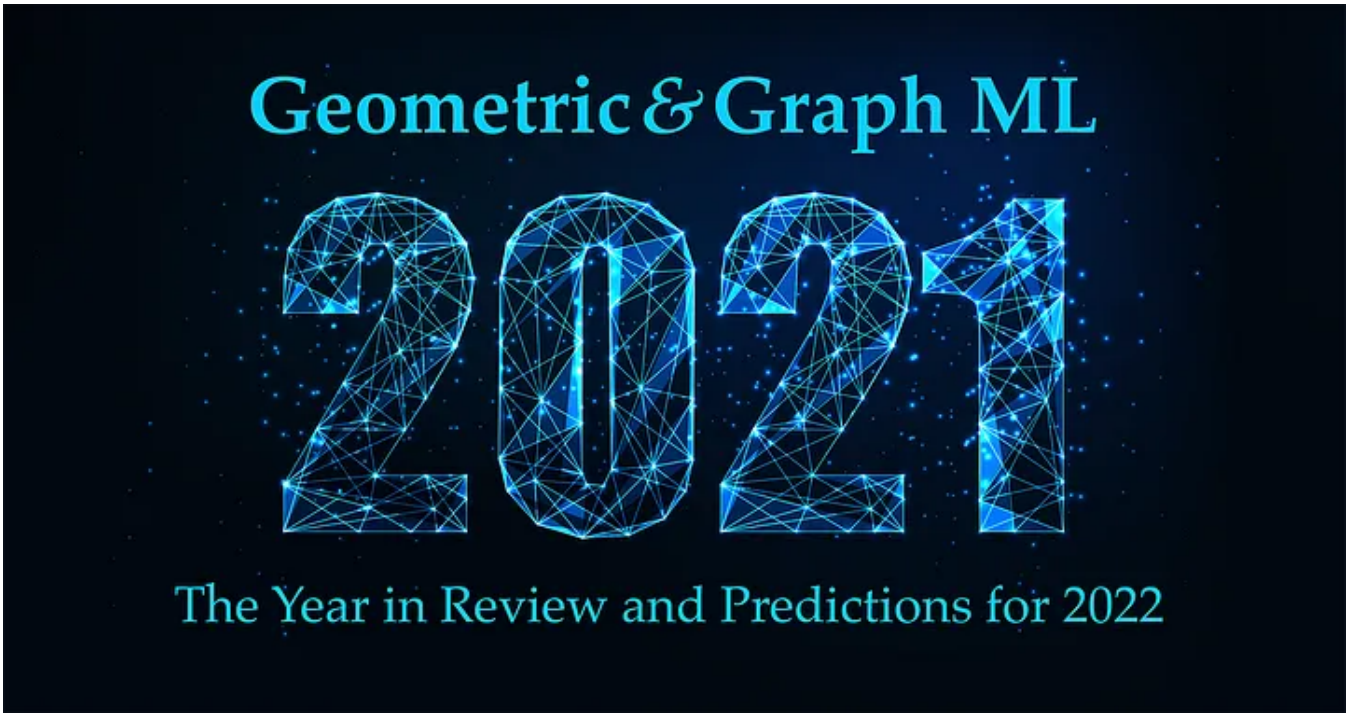


Image: Shutterstock

This post was co-authored with Petar Veličković. See also my last year's prediction, Michael Galkin's excellent post on the current state of affairs in Graph ML, a deeper dive into subgraph GNNs, techniques inspired by PDEs and differential geometry and algebraic topology, and how the concepts of symmetry and invariance form the picture of modern deep learning.

Summing up impressions of 2021 and forecasting the year ahead with the help of the leading domain experts is a worthwhile and rewarding experiment, as it gives a range of diverse insights and opinions to learn from. In the words of one of our interviewees, who keeps an attentive eye on the top emerging trends across the entire machine learning landscape:

> "With geometric and graph-based ML methods going from niche to one of the hottest fields of AI research today, it is exciting to imagine what the next twelve months will hold." — Nathan Benaich, General Partner of Air Street Capital and co-author of the State of AI Report

In retrospect, this effort turned out not as easy a feat as we have initially thought, since the field has hugely grown in the past year — this is witnessed by the length of this post, the longest we have written so far.

Quick take-home messages (if you are lazy to read the rest)

1. **Geometry becomes increasingly important in ML.** Differential geometry and cognate fields have brought new ideas including new equivariant GNN architectures that exploit symmetries and graph analogues of curvature as well as understanding and utilising uncertainty in deep learning models.

2. **Message passing is still the dominant paradigm in GNNs.** In 2020, the community came to terms with the limitations of message-passing GNNs and looked beyond this paradigm for more expressive architectures. In 2021, it is evident message passing still has several aces up its sleeve, as several works

showed the possibility to achieve better expressive power applying GNNs to subgraphs.

3. **Differential equations give rise to new GNN architectures.** The trend that started with NeuralODEs expanded to graphs. Several works showed how to formulate GNN models as discretisations of continuous differential equations. In the immediate term, this will lead to new architectures that avoid the common plights of GNNs such as oversmoothing and oversquashing. In the long run, we will likely gain a better understanding of how GNNs work and how to make them more expressive and explainable.

4. **Old ideas from Signal Processing, Neuroscience, and Physics get a new life.** Graph signal processing is considered by many the spark that re-ignited the recent interest in Graph ML and provided the first set of tools such as generalised Fourier transforms and graph convolutions. Other fundamental techniques on which classical Signal Processing and Physics stand, such as Representation Theory, have led to some important advances in 2021 and still have much in store.

5. **Modeling complex systems requires going beyond graphs.** The 2021 Nobel Prize in Physics was awarded to Giorgio Parisi for the study of complex systems. While at the basic level of abstraction such systems can often be described as graphs, one has to account for more complex structures such as non-pairwise relations and dynamic behaviours. Multiple works in 2021 addressed dynamic relational systems and showed how to extend GNNs to higher-order structures such as cell- and simplicial complexes that are traditionally dealt with in the field of algebraic topology. We will likely see other even more exotic objects from this field adopted in ML.

6. **Reasoning, axiomatisation, and generalisation are still big open questions in Graph ML.** Throughout the year, we have seen continued advances in GNN architectures inspired by algorithmic reasoning, as well as more robust out-of-distribution generalisation on graph-structured tasks. We now have Knowledge Graph reasoners that explicitly align with the generalised Bellman-Ford algorithm, as well as graph classifiers that leverage an explicit causal model of distribution shifts. Arguably, directions such as these uncover a goldmine of future directions for more robust and general GNNs; it is likely that many of these will be explored thoroughly in 2022.

7. **Graphs become increasingly popular in Reinforcement Learning, but probably still have a way to go.** Perhaps unsurprisingly, Reinforcement Learning is full of graphs and symmetries: often either in the structure of the RL agent, or in representations of the environment. 2021 has seen several research directions that attempt to exploit this structure, to varying levels of success. We now have a better understanding of how to exploit many of these symmetries in RL, including in multi-agent systems; however, modelling agents as graphs appears to not require utilising the graph structure strictly. Despite such mixed feelings, we believe 2022 is a year that holds a lot of promise for graph- and geometry-powered RL.

8. **AlphaFold 2 is a triumph of Geometric ML and a paradigm shift in structural biology.** The possibility to predict the 3D folding structure of a protein was hypothesised in the 1970s by the Chemistry Nobel Prize winner Christian Anfinsen. It turned out a very hard computational task, a Holy Grail of structural biology. In 2021, DeepMind's AlphaFold 2 broke the record of the previous version, achieving accuracy that convinced domain experts and spurred broad adoption. At its core, AlphaFold 2 is a geometric architecture based on equivariant attention.

9. **Drug discovery and design benefits from GNNs and their confluence with Transformers.** Few in the ML community are not aware that the origins of GNNs can be traced to works in computational chemistry from the 1990s. It is thus not surprising that the analysis of molecular graphs is among the most popular applications of GNNs. In 2021, there has been continuous significant progress in this field, with dozens of new architectures and several benchmark-beating results. One of the winners was an adaptation of Transformers to graphs, that promises to emulate the keys for the success of these architectures in NLP: large pre-trained models capable of generalising across tasks.

10. **AI-first drug discovery is increasingly using Geometric and Graph ML**. The success of AlphaFold 2 and molecular GNNs bring closer the dream of new AI-designed drugs. Alphabet's new company Isomorphic Labs are an indication of industry's betting on this technology. However, modelling interactions of molecules is the important frontier that has to be crossed in order for these hopes to materialise.

11. **Quantum ML benefits from graph-based methods.** Quantum ML is still an exotic niche for the majority of experts in the field, but quickly becoming a

reality with the availability of quantum computing hardware. Recent works from Alphabet X showed the advantages of graph structural inductive biases in Quantum ML architectures, marrying these apparently unrelated fields. In the long run, geometry might play a more fundamental role, as quantum physical systems often possess rich and esoteric group symmetries that could be leveraged for quantum architecture design.

Geometric and Graph-based ML methods shone in high-profile applications in 2021. Images: Nature/Science.

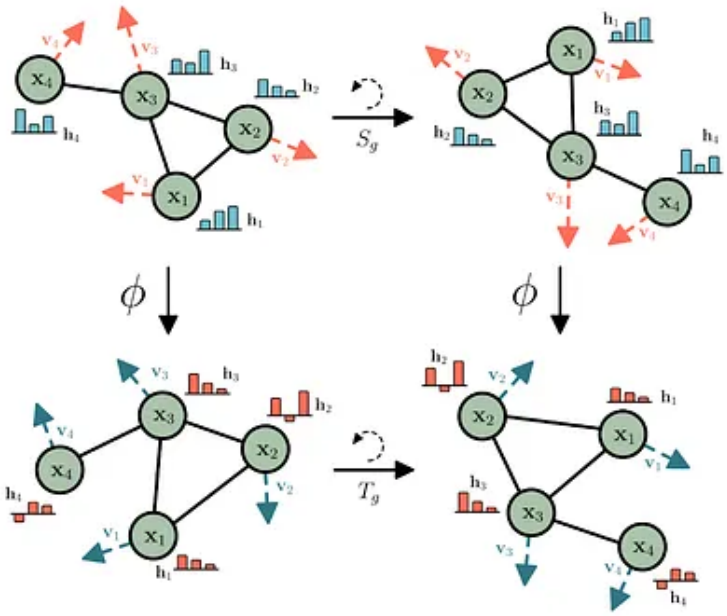
Geometry becomes increasingly important in ML

If we had to choose one word that continuously permeated virtually every area of graph representation learning in 2021, there is little doubt that *geometry* would be a prime candidate [1]. We wrote about this last year, and our interviewees definitely seem to agree — more than half of them called out this keyword in one way or another.

> "In the past year, we have seen the use of many classical geometric ideas in new ways in Graph ML." — Melanie Weber, Hooke Research Fellow, Mathematical Institute at the University of Oxford

Melanie adds: "Notable examples include utilising symmetries to learn models more efficiently, the application of ideas from optimal transport, or the use of curvature notions from differential geometry in representation learning.

Recently, there has been a surge of interest in understanding the geometry of relational data and utilising those insights to learn good (Euclidean or non-Euclidean) representations [1]. This has given rise to many GNN architectures that encode specific geometries. Notable examples are Hyperbolic GNN models [2], which were first introduced in late 2019 as tools for learning efficient representations of hierarchical data. In the past year, we have seen a surge of new models and architectures that learn hyperbolic representations more efficiently or are able to capture more complex geometric features [3–4]. Another line of work utilises geometric information in the form of equivariances and symmetries [5]."



This year we saw a surge of geometric techniques in the context of graph neural networks, such as equivariant message passing that prove itself crucial in biochemical applications from small molecule property prediction to protein folding. Image from Satorras et al. [5].

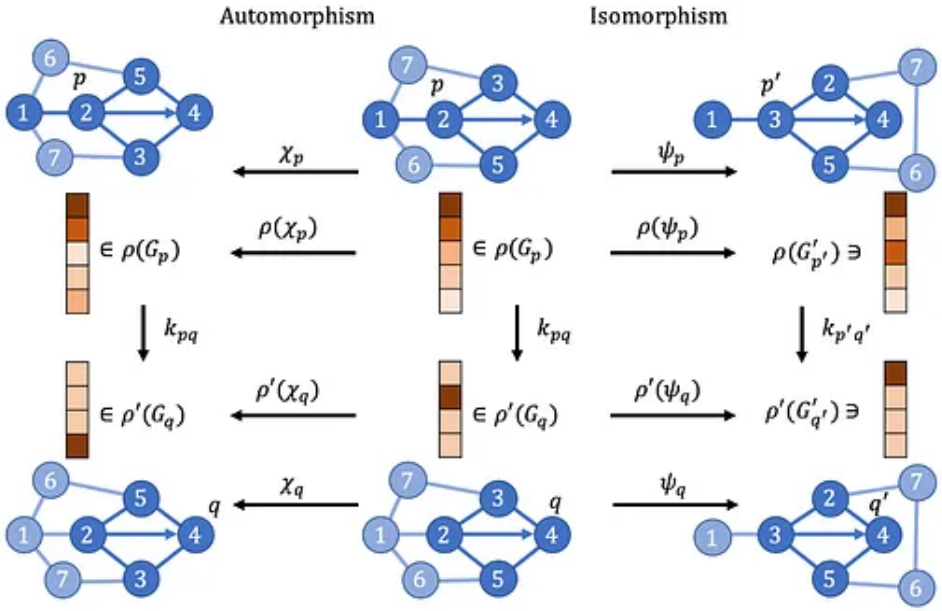
Melanie doubled down on differential geometry, indicating its many prospective applications for 2022: "Discrete Differential Geometry, which studies the geometry of discrete structures like graphs or simplicial complexes, has been used to gain insight into GNNs. Discrete notions of curvature are important tools for characterising local and global geometric properties of discrete structures. A notable application of curvature in Graph ML was recently proposed in [6], where discrete Ricci curvature is studied in the context of graph rewiring, proposing a new approach for mitigating over-squashing effects in GNNs. It seems likely that discrete curvatures will be linked to other structural and topological questions in Graph ML in the future.

I expect that those themes will continue to influence the field in 2022 and also make their way into more applied areas of Graph ML. This may drive advances on the computational side with the aim of mitigating computational challenges related to the implementation of non-Euclidean algorithms with tools that are traditionally designed with Euclidean data in mind. In addition, geometric tools, such as discrete curvatures, can be expensive to compute, making it difficult to integrate them into

We also heard an abridged account of the many new superpowers GNNs gained in recent years by virtue of this geometric lens:

> "Graph neural network designers have been increasingly recognising the rich symmetry structure of graphs." — Pim de Haan, PhD student, University of Amsterdam

"GNNs traditionally performed permutation-invariant message passing and later used Group- and Representation Theory to construct equivariant maps between representations of the permutation group of nodes [7]. Recently, in close analogy to the local symmetries of manifolds (called 'gauge symmetries' [8]), we have started studying the local symmetries in graphs arising from isomorphic subgraphs. We found that these form not a group, but a category of symmetries [9], and that incorporating these into neural network architectures improves the performance of graph ML tasks such as molecular prediction [10]", outlines Pim.



Graph ML researchers exploited richer symmetry structures of graphs. Image from de Haan et al. [9].

Inspired by these important findings, Pim offers his prognosis for 2022: "For the new year, I am hoping to see Category Theory become widespread as a design language for neural networks. This would give us a formal language to discuss and exploit more sophisticated symmetries than before. In particular, I am excited to see it used to tackle local and approximate symmetry of graphs, combine the geometric and combinatorial structure of point-clouds, and help us study the symmetries in causal graphs."

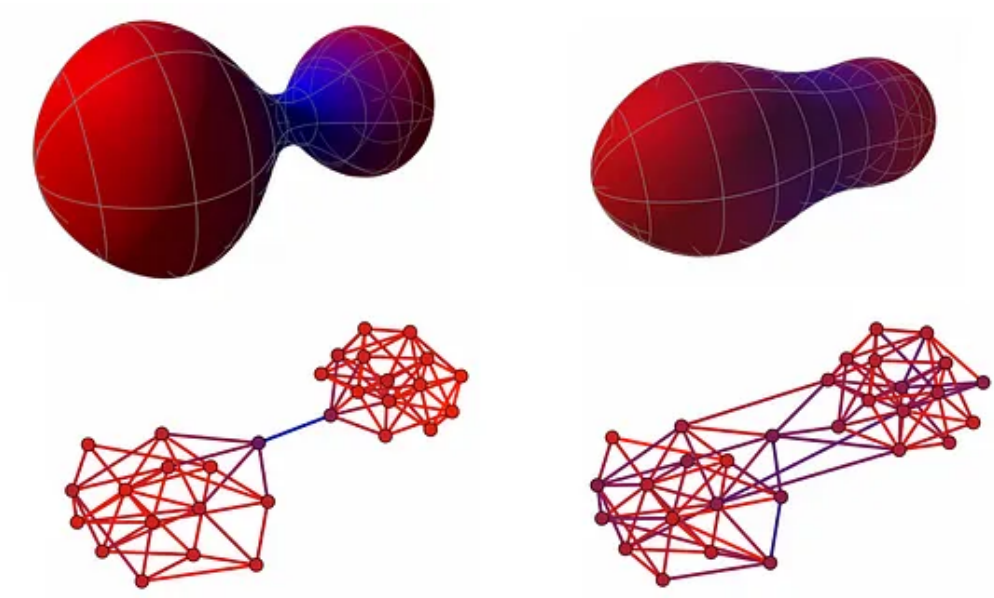
By now, it should be clear that GNNs stand to gain a lot if they embrace the fact most graphs originate, or could be interpreted, as discretisations of some continuous manifolds. We solicited expert opinion on this from a differential geometer turned machine learner — here is what he had to say:

> "Although graphs are not differentiable, many ideas that have been successfully applied in the analysis of manifolds are now surfacing in the GNN realm." — Francesco Di Giovanni, ML Researcher at Twitter

Francesco recounts and forecasts some of the key work in the area: "I am particularly interested in PDE methods that had first been borrowed from the study of surfaces to process images [11–12]). The latter has also explored ideas of graph rewiring — a term referring to modifications of the underlying adjacency — that fall

within the larger spectrum of geometric-flow approaches where given objects are suitably modified to aid classification. This principle has also guided our recent work [6], where we leveraged a new notion of edge-based curvature to investigate the problem of over-squashing in GNNs (first observed by Uri Alon [13]) and propose a graph rewiring method aimed at removing bottlenecks that lead to such issue. The role of geometry in the form of symmetries preservation and breaking is also now arguably a key factor in the application of GNNs to molecules, as for example in [14].

I believe that we are only moving the first steps into this new direction. Graph-rewiring techniques as above will potentially play a role in addressing some of the main limitations of message passing both in the form of performance on heterophilic datasets and tackling long-range dependencies. I also hope that we will soon bridge a somewhat conceptual gap between convolution on graphs and convolution on manifolds, which could lead to the next generation of GNNs. Finally, I would be happy to see geometric variational methods shedding some further light on the intrinsic dynamics of GNNs and hopefully giving rise to more principled ways to design new GNNs architectures and compare existing ones."



Concepts from differential geometry such as Ricci curvature and geometric flows were used in the Graph ML context to improve information flow in GNNs. Image from J. Topping et al. [6].

In the broader context of machine learning, differential geometry has more in store than just improving GNNs:

"Differential geometry and its mathematical siblings are often used in the hope of providing well-founded

> solutions to problems whose precise formulation creates nonlinear geometries." — Aasa Feragen, Assistant Professor, University of Copenhagen

Aasa outlines, in particular, her enthusiasm for differential geometry in estimating uncertainties: "One lesson from 2021 is that differential geometry plays a fundamental role in understanding and utilising uncertainty in deep learning models. A great example is using model uncertainty to yield geometric representations of data that reveal biological information which remains obscured in standard Euclidean representations [15]. Another example is leveraging the Riemannian geometry encoded by local directional data to allow for a natural quantification of uncertainty in estimated structural brain connectivity [16]."

Ultimately, Aasa hopes that this area will grow in 2022: "I had expected 2021 to address uncertainty quantification within geometric machine learning far more widely. Geometric models are often applied to data that is heavily preprocessed in order to reveal its geometric structure. In other words, our data is often estimated from raw data, which comes with errors and uncertainties. I think it is high time that we start to assess the effect of these raw data uncertainties on the objects we treat as data, as well as how that uncertainty should propagate to our models. This is my main hope for 2022, that we turn to properly incorporating measurement errors into our analysis of non-euclidean data. This would probably gain more traction if we — as scientists and representatives of our own scientific disciplines — made an effort to tear down the community divide between statistics and deep learning."

Message passing is still the dominant paradigm in GNNs

As the field of Graph ML came to terms with the "fundamental limitations of the message-passing paradigm" (to quote Will Hamilton) due to its equivalence to the Weisfeiler-Lehman isomorphism test, Michael's 2021 post predicted that "progress will require breaking away from the message-passing schemes that dominated the field in 2020 and before." This forecast materalised only partially: while 2021 has brought multiple works on more expressive GNN architectures, most of them still remain largely within the remits of the message passing paradigm.

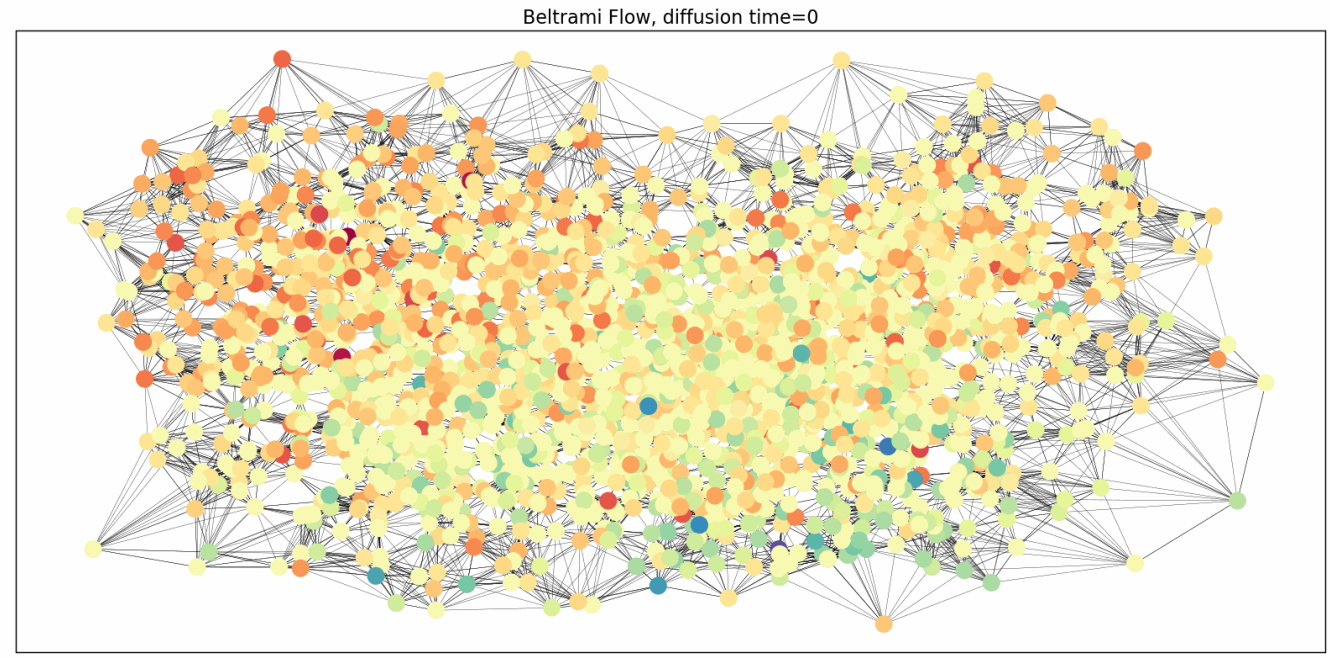
One such example is the recent line of concurrent works on the use of subgraphs to improve the expressive power of GNNs. Haggai Maron, who also provided his input

in last year's post, explains:

"The idea behind "Subgraph GNNs" is to represent graphs as collections of their substructures, a theme which can be found in the famous works by Kelly and Ulam in the 1960s on the Graph Reconstruction Conjecture [17]. Nowadays, the same idea is used to construct expressive GNNs which, in turn, raise new, more refined Reconstruction Conjectures."

> "I expect subgraph GNNs and the corresponding reconstruction conjectures to be a fruitful research direction in the years to come. "— Haggai Maron, Research Scientist, Nvidia

Differential equations give rise to new GNN architectures



Several approaches for deriving Graph Neural Networks from discretised diffusion-type PDEs were shown in 2021. Image by James Rowbottom based on [12].

Another trend in 2021 was reformulating learning on graphs through the dynamics of physical systems, expressed as differential equations. Same way as Ordinary Differential Equations served as a powerful tool to understand residual neural networks ("Neural ODEs" were crowned as the best paper at NeurIPS 2019), Partial Differential Equations can model information propagation on graphs and allow to recover many standard GNN architectures as iterative numerical schemes solving

such PDEs [18]. In this formulation, the graph is regarded as a mere *discretisation* of a continuous object, and in the words of our correspondent:

> "This brings an alternative and refreshing view on how GNNs can be used to extract meaningful information for downstream ML tasks and moves the focus away from the domain that supports information and instead uses graphs as a support for computations over signals" — Pierre Vandergheynst, Professor at EPFL

Pierre sees this as a broader direction for the coming year: "In 2022, I expect that this will become a new trend: using graphs as a mechanism to perform localised coherent computations, exchange information about said computations over the dataset and use that as a mechanism to zero in on the overall properties of the data. That would be an exciting avenue for unsupervised or zero-shot learning."

Old ideas from Signal Processing, Neuroscience, and Physics get a new life

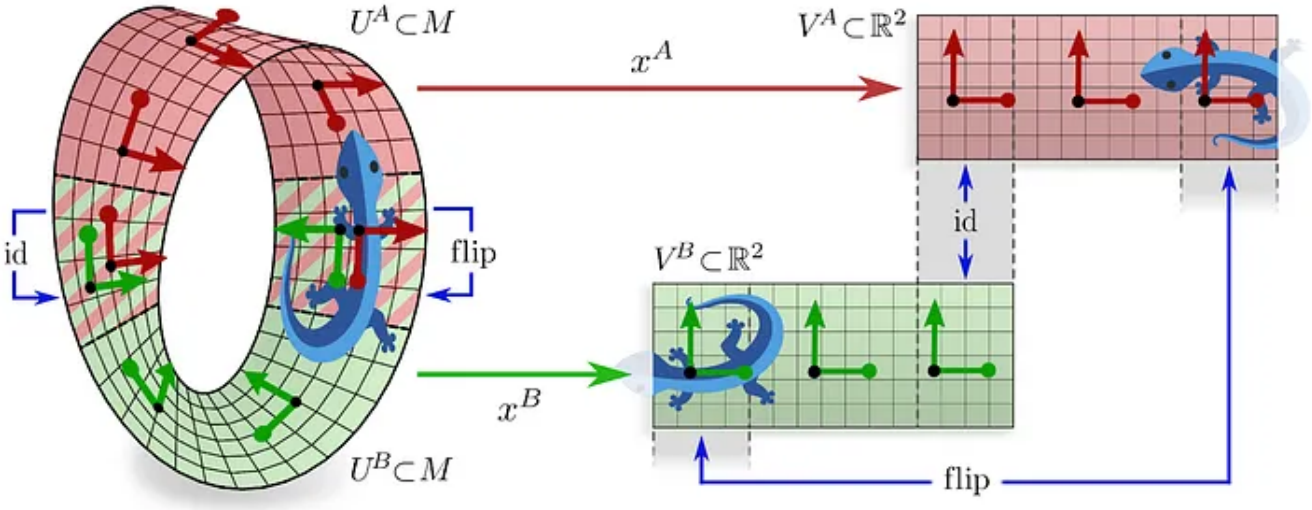
Many of the modern GNNs have their roots in methods originally developed in the signal processing community. Pierre Vandergheynst, one of the fathers of Graph Signal Processing (GSP) [19], offered an interesting perspective on the progress in Graph ML methods from this perspective:

"Graph Signal Processing started with two directions to enrich digital signal processing. The first was to generalise the domain that supports information, moving from the traditional setting of low-dimensional Euclidean spaces and allowing signals to be defined on much more complex but structured objects that can be represented by graphs (networks, meshed surfaces, etc.). The second motivation for GSP was to move away from structured domains and directly work on some dataset, using a graph (a nearest neighbour of some sort) to represent similarities between samples. The underlying idea was that label fields inherit some regularity that can be defined over the graph and captured by appropriate transforms. In this sense the graph becomes the support of localised computations over the entire dataset. Some of the interesting ideas in GNNs are rooted in these

early motivations and it is interesting that 2021 produced highlights that continue this trend."

> "One of the appeals of classical linear transforms (e.g. Fourier or wavelet) is that they provide a universal "latent" space that comes with mathematical properties: e.g. smooth signals have low-frequency Fourier coefficients, while piecewise-smooth signals have sparse and localised wavelet coefficients" — Pierre Vandergheynst, Professor at EPFL

Pierre further recollects: "There used to be a whole tradition of constructing linear transforms to reveal signal properties. Physicists in particular lead the way in designing equivariant transforms for different symmetries based on group actions, e.g. the continuous wavelet transform for the affine group (or the similitude group in higher dimensions), linear time-frequency analysis for the Weyl-Heisenberg group, and so on. The literature on the (somewhat arcane) domain of Coherent States in mathematical physics offers a generic recipe: construct a linear transform by parametrising a system of functions using group representations. Add a non-linearity and learnable parametric functions, and you get close to some of the nicest papers of 2021 endowing GNNs with symmetries that make them super useful for problems in physics or chemistry [20]."



Group representations, a tool that has seen traditional use in signal processing and physics, allow deriving coordinate-free deep learning architectures that can be applied to manifolds. Image from Weiler et al. [20].

Pierre expects that the trend of building structured latent spaces will continue in 2022, "driven in part by the need of applications but also by the desire to trade adaptivity and explainability: structured transform domains are not adaptive but very readable and GNNs might help strike a good balance."

One area of science that traditionally goes intimately together with signal processing is neuroscience; indeed, much of what we know about how animals perceive the world around them is by analysing the electrical signals transmitted by the brain. We caught up with a prominent researcher at the intersection of machine learning and neuroscience–and she readily states the importance of graphs in this particular domain:

> "My background is in Computational Neuroscience, and I first got into graphs because I wanted to represent how humans and animals learn structure."
> — Kim Stachenfeld, Research Scientist at DeepMind

"Graphs are handy mathematical objects for capturing how humans and animals might represent related concepts acquired across isolated fragments of experience, and stitch it together into a globally coherent, integrated body of knowledge.", Kim explains.

Kim shares the excitement others have highlighted about geometry, while also emphasising the links between graph signal processing and relevant neuroscience research:

"An area of research I'm really excited about that has seen a lot of action this year is work that combines localised operations in neural networks with representations of underlying or intrinsic geometry. Prominent examples include some of the new works this year on equivariance in GNNs [9,5], which allow GNN models to leverage geometry and symmetries external to the graph structure.

Another work in this vein uses Graph Laplacian eigenvectors as position encoding for graph transformers, allowing the GNN access to information about the intrinsic,

low-dimensional geometry without being overly constrained by it [21]. In addition to having deep mathematical roots and having been part of the GNN literature for a long time [22–23], they have connections to research into the neuroscience of geometric, relational reasoning [24–26]."

Beyond all this, Kim is generally openly excited about the various applications GNNs have seen in 2021, both within neuroscience and more broadly:

"I am super excited about the increasing number of diverse applications of GNNs, particularly to very large-scale, real-world data. Examples include using GNNs to predict traffic outcomes [27], to simulate complex physical dynamics [28], and to solve problems over particularly large-scale graphs [29]. I'm also excited to see a growing interest in applying GNNs to the analysis of neural data [30–32]. These problems have real-world impact, and they challenge our models to scale efficiently and generalise while still capturing really complicated dynamics — exactly the balance of structure and expressivity GNNs are optimised for. "

Modelling complex systems requires going beyond graphs

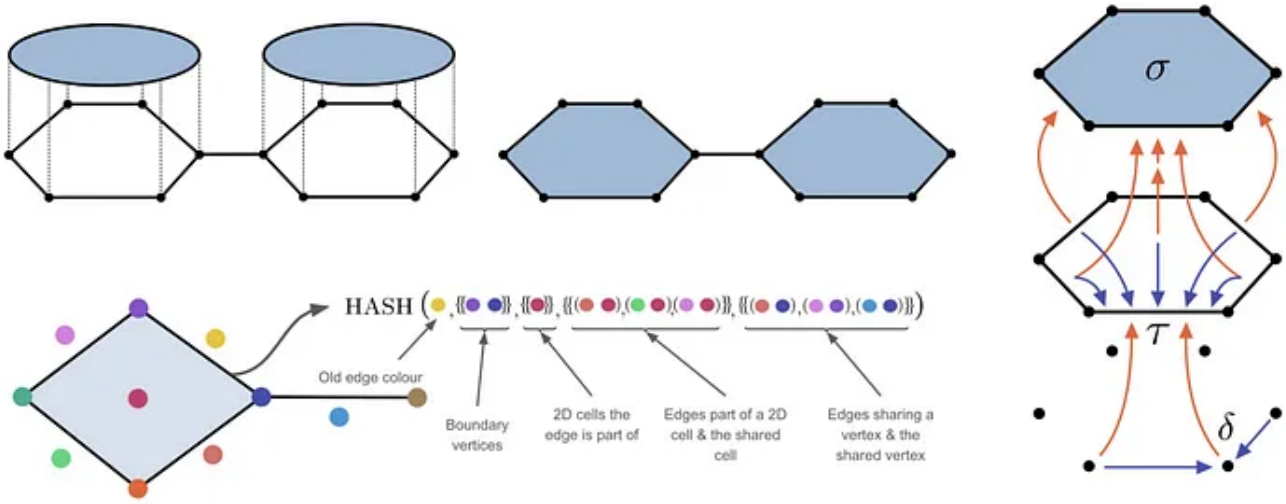
> "The theme of the 2021 Nobel Prize in Physics was the study of complex systems. At the most fundamental level, complex systems consist of entities and their interactions. Complex systems are often represented as complex networks, which fuel the work in graph ML." — Tina Eliassi-Rad, Professor, Northeastern University

"As graph ML comes of age, we need to scrutinise the system dependencies that can manifest themselves in different flavors (subset, temporal, and spatial), the common mathematical representations (graphs, simplicial complexes, and hypergraphs), their underlying assumptions, and the dependencies they encode [33]. The choice of mathematical representation is important because information can be lost (or imputed) when we convert data from one representation to another.

There is no perfect way to represent a complex system, and modelling decisions made when examining a data set from one system are not necessarily transferable

to another system, or even to another data set from the same system. Nevertheless, exciting research opportunities open up for graph ML when we consider system dependencies in relation to the mathematical representations we choose [33–36].”

Pierre Vandergheynst highlights that graphs do not always offer an adequate model of complex systems, and one may have to go beyond graphs: “several beautiful papers have brought new structured information domains that are captured by generalisations of graphs. A striking example is the use of simplicial complexes and other ideas of algebraic topology to construct new neural networks that experimentally and provably improve on GNNs [37]. This is a trend that will surely continue in 2022 as we learn to dig into a vast array of structured mathematical objects offered by algebraic topology or differential geometry.”



Lifting graphs to cellular or simplicial complexes allows more complex topological message passing that yields GNN architectures that go beyond the Weisfeiler-Lehman test expressive power. Image from Bodnar et al. [38].

Cristian Bodnar is equally enthusiastic about the connections between algebraic topology and Graph ML: “In the past year, convolutional [39–41] and message passing [37–38,42–43] models on simplicial and cell complexes addressed many of the limitations of GNNs such as detecting certain substructures, capturing long-range and higher-order interactions, handling higher-order features and escaping the WL hierarchy. In practice, they obtain state-of-the-art empirical results in molecular problems [38] and trajectory prediction and classification [37,41] tasks. In 2022, I expect these approaches to expand to exciting new applications such as hard problems in computational algebraic topology [44], link prediction [45], and computer graphics [46].”

Like Pierre, he thinks that algebraic topology still has a lot more to offer to machine learning:

> "We will likely see the adoption of even more exotic mathematical objects, which have remained relatively unexplored so far, such as cellular sheaves [47–48] and quivers [49]. At the same time, I believe that these topological approaches will provide a novel set of mathematical tools for analysing and understanding GNNs."– Cristian Bodnar, PhD Student, Cambridge University

Another very important example of a complex networked system are spatio-temporal graphs, whose structure evolves over time. We solicited the opinion of a prolific researcher in the area on her thoughts and hopes for 2022:

> "I am very excited about the role of Graph ML in learning spatio-temporal dynamics." — Rose Yu, Assistant Professor, UC San Diego

Rose elaborates: "Applications in COVID-19 prediction [50-Wu et al 2021], traffic forecasting [51-Shang et al 2021] and trajectory modelling [52-Walters et al 2021] require capturing complex dynamics of highly structured time series data. Graph ML provides the capability to capture the spatial dependency, the interactions between time series, and the correlations in the dynamics.

In 2022, I would be happy to see the blending of ideas from time series and dynamical systems with Graph ML. Hopefully, these ideas would generate new model designs, training algorithms, and rigorous understandings of the internal mechanism of complex dynamical systems."

Further, Rose believes that symmetry discovery is an important overlooked open problem in graph representation learning:

"Graph neural networks contain permutation symmetry (invariance or equivariance). But this global symmetry can be fundamentally limiting. There are many excellent papers generalising graph neural networks to symmetry groups beyond permutation and to local symmetries (e.g. [9,5,52].

In 2022, I would be interested to see more investigation into the symmetries in graph neural networks. For example, the new development in equivariant networks including automatic symmetry discovery [53] can be translated to Graph ML."

Reasoning, axiomatisation, and generalisation are still big open questions in Graph ML

The development of entirely novel inductive biases for GNN layers has generally slowed down in recent years, as increasingly many standard GNN benchmarks get saturated. Indeed, even some of the tasks offered by the Open Graph Benchmark now seem to rely on non-architectural tricks (such as external data or output post-processing) to advance.

A notable exception to the above comes from reasoning problems; therein, edges are more than a simple guide for aggregation, rather, they are a recipe for how to pass data in the graph. Further, for true reasoning capability we also often require a degree of generalisation that is rarely required in graph tasks, especially transductive ones. Clearly, off-the-shelf message-passing models are going to struggle here; therefore, both reasoning and its close cousin, simulation, have been responsible for some of the most interesting novel GNN updates in 2021.

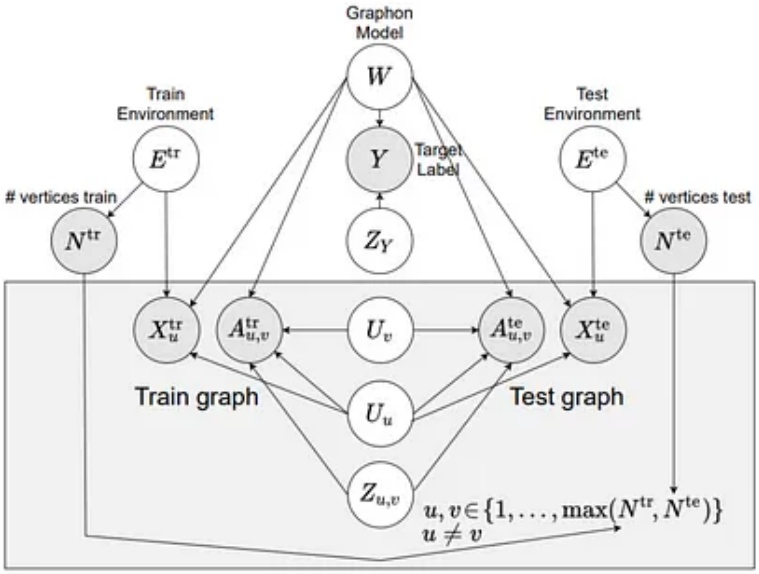
Because the field is so nascent, there are many ways to go about encouraging generalisation; especially out-of-distribution. We got in touch with one of the proponents of a very exciting direction proposed in 2021. His response starts by sending us back to the past:

> "My attempt to peer into the future (2022) must start with a dive into the distant past and its connections to 2021 papers." — Bruno Ribeiro, Assistant Professor, Purdue University

"Karl Pearson, R. A. Fisher, and Charles Spearman (some of the founders of modern statistics) had two well-documented blind spots: (i) Causality (Pearson famously said

something akin to correlation is all you need for causation); and (ii) Axiomatic probability (developed in parallel by Kolmogorov in Russia). Their blind spots have interesting ramifications in Graph ML today."

Causality has emerged as an exciting view into generalisation with GNNs in 2021. Bruno elucidates how: "Causality is a blind spot for both ML and Graph ML. In 2021, there were a few Graph ML works in this space. One of the simplest applications of causality addresses out-of-distribution (OOD) tasks. In ML deployments, one can see test and training data with different distributions [54–55]. Being robust to these distribution shifts, without access to test distribution data, requires the classifier to be invariant to the shift. This invariance is now known as counterfactual invariance [56–58]. For instance, [56,59] show that while GNNs can create representations of any-size graphs, graph classification exclusively trained on small (large) graphs generally performs poorly on large (small) test graphs (an OOD task). Focused on counterfactual invariance, [56] uses Lovász & Szegedy graph limit results [60] to show that for a family of graphs (graphons), one can use induced subgraph representations to obtain a size-invariant whole-graph representation for the above OOD task. In a different line, [61] proposes to learn graph representations based on task-relevant subgraphs that are invariant to spurious (task-irrelevant) subgraphs, which allows the graph classifier to survive spurious OOD subgraph distribution shifts in such tasks. I expect 2022 will bring new exciting connections between Graph ML and causality."



The work of Bevilacqua et al. [56] has demonstrated how size-invariant OOD generalisation can be achieved for graph classification tasks, with the use of an explicit causal model (depicted here) that models this distribution shift.

Beyond causality and its implications to OOD, Bruno also highlights the strides made in underpinning GNN architectures with careful axiomatisation: "Axiomatic probability works by stating properties of the distribution that generated the data rather than giving it a functional form. For instance, we can say the distribution of a sequence of random variables is permutation invariant (exchangeable) and this is all we know about the data. Pearson, Fisher, and Spearman likely would be bewildered that one can create a predictive model assuming only permutation invariance (or equivariance)."

"Indeed, between 2013–2020, the Graph ML community learned to partially fill this blind spot, which paved the way to the rapid growth of Graph ML inside ML. In practice, this geometric axiomatization is not yet complete. State-of-the-art results still require building neural architectures whose designs are not yet fully driven by known data properties (axioms), and it is no surprise that SOTA now is partially driven by data post-processing (to conform the data to the architecture). But 2021 saw some progress towards further geometric axiomatisation: graph representation architectures were shown to be aided by extra axioms (e.g. related to induced subgraphs and the reconstruction conjecture [62–66]) and can now learn to avoid information bottlenecks using the graph's topological curvature [6].

In 2022, we will hopefully see further developments in the axiomatization of Graph ML architectures. An interesting direction is the axiomatization of link prediction to fill a gap left by Spearman, as foundationally described as joint structural representations in 2020 by [67] and furthered with methods in 2021 by [68–69] and others, which in 2022 may play a role in the pursuit of causal link prediction methods (say, for recommender systems [70])."

Another area of machine learning that stands a lot to gain from reasoning and generalisation is knowledge graph reasoning. KGs have been around since the early days of artificial intelligence, and given their large industrial implications, they are certainly here to stay. In recent work, Jian Tang has leveraged direct knowledge from pathfinding algorithms to construct a state-of-the-art KG inference method. Jian recounts this project in his response to us:

"We observed significant progress in knowledge graph reasoning in 2021. Existing knowledge graph reasoning algorithms usually only work in transductive settings but not inductive settings. In last year's NeurIPS, we proposed a general and flexible framework based on GNNs for link prediction on graphs, called Neural Bellman-

Ford Networks (NBFNet) [71]. The NBFNet is very general, covers many traditional path-based methods, and can be applied to both homogeneous graphs and multi-relational graphs (e.g., knowledge graphs) in both transductive and inductive settings. It outperforms existing methods by a large margin in both transductive and inductive settings, achieving new state-of-the-art results. I believe NBFNet will become the standard framework for link prediction on graphs in the future."

Graphs become increasingly popular in Reinforcement Learning, but probably still have a way to go

Reinforcement Learning remains one of the most prominent and active areas of AI research today. Conveniently, and perhaps unsurprisingly, it is full of graphs and symmetries: often either in the structure of the RL agent, or in representations of the environment. But even the central mathematical objects in RL, Markov Decision Processes (MDP), are effectively graphs of states and actions. Last year, Petar predicted that in 2021 we would see a sweeping performance of GNNs in RL. Did this prediction become reality?

Surprisingly, our two correspondent practitioners of geometric methods in RL had diametrically opposing conclusions on this:

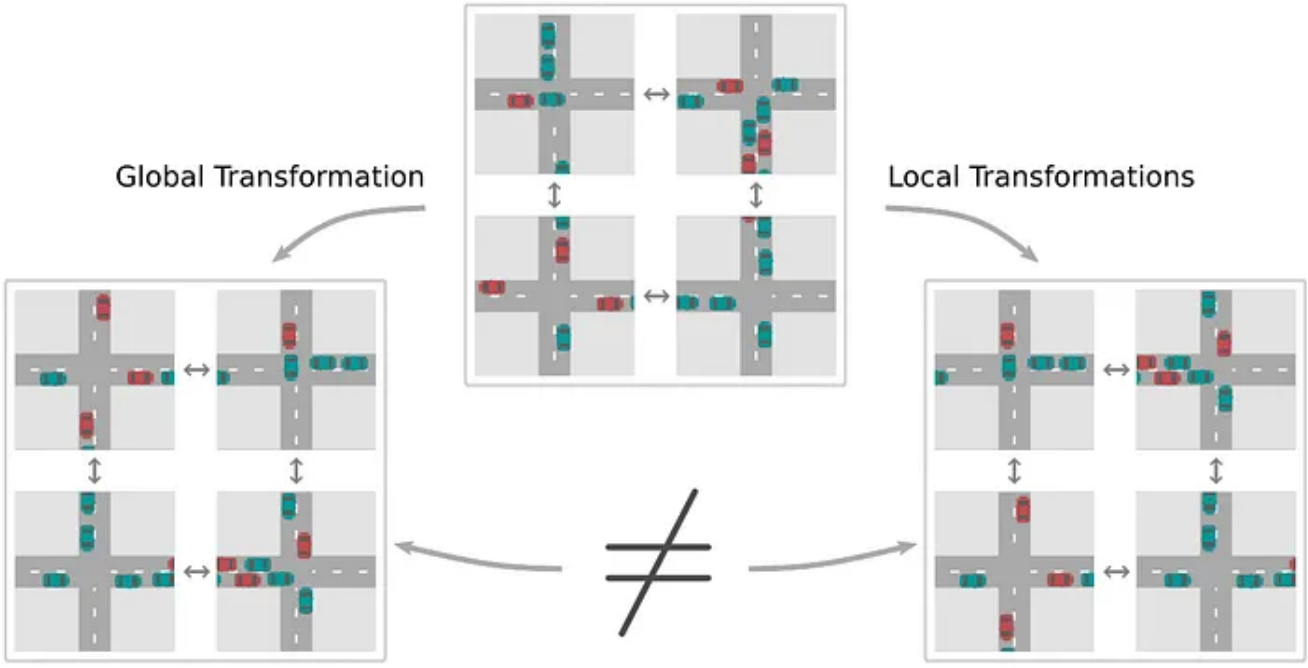
> "In 2021, graphs became full-fledged players in RL." — Vitaly Kurin, PhD Student, University of Oxford

> "Graph Neural Networks have made a big impact on many areas of machine learning. A notable exception is Reinforcement Learning, where in 2021, vanilla Graph Neural Networks have still not made a huge splash, despite initial optimism a few years ago [72–74]." — Elise van der Pol, PhD Student, University of Amsterdam

While the conclusions differ, the body of work highlighted by both is largely similar.

Vitaly highlights: “In 2021, graphs appeared in RL through two noticeable groups. The first uses graph representations (including attention-based models) for established RL benchmarks to improve generalization/transfer abilities. Examples include our work in continuous control [75–76], multi-agent RL research [77–78] , and robot co-adaptation [79–80]. The second group, which I am very excited about, expands the RL application area to environments with graph-based state and action spaces. This group includes power grid management [81], combinatorial optimization [82], and finding counter-examples to open conjectures in graph theory [83].”

Elise remarks: “The work of Kurin et al. [75] presents empirical results showing that vanilla graph networks underperform in reinforcement learning compared to transformer-based methods even when the underlying environments are graph-structured. While graph networks have not become a more integral part of reinforcement learning yet, several works from 2021 highlight specific scenarios where reinforcement learning can benefit from graph networks. Highlights include the combination of implicit planning with self-supervised learning [84], using GNNs to include inductive biases [85], and applications such as network planning [86] and compiler optimization [87]. Additionally, in 2021 we also saw some first investigations into Transformer [8] methods that exploit the sequential structure of reinforcement learning tasks [89–90]. There has also been work on using Transformers and GAT [91] methods in multi-agent reinforcement learning [92–93]. Additionally, we saw more integration of GNNs with more types of equivariance, for example in our work on equivariant multi-agent coordination [94].”



2021 saw exciting applications of Geometric Deep Learning to multi-agent reinforcement learning systems. It enabled the support of policy functions that are equivariant to global symmetries of the task, while still supporting distributed execution. Image from van der Pol et al. [94].

Despite their differences in summarising 2021, both are optimistic about how Graph Machine Learning will shape RL in the year ahead!

Vitaly claims: “To me, graphs are key in bringing RL to the real world since they are omnipresent both in industrial (chip design, code modelling, drug discovery) and scientific applications (physics, chemistry, biology, economics). I believe 2022 will witness a Cambrian explosion of novel RL benchmarks that will channel RL research in a more applied direction and motivate fundamental research while revealing inconsistencies in our thinking about RL.”

“My expectation for 2022 is that transformers-based and GAT-like methods will become more prominent in reinforcement learning, given their initial success over vanilla graph networks. There is also a strong potential along the veins of combinatorial optimization with graph networks, equivariance, and reinforcement learning, for example in molecular design, network planning, and chip design. Lastly, there is room for more interplay between bisimulations and MDP homomorphisms and GNN-based models, especially when it comes to discovering the abstract graphs underlying (factored) reinforcement learning problems, as well as learning to partition MDPs in a principled manner. Will 2022 become the year where graph networks become a go-to component in reinforcement learning?”, questions Elise.

AlphaFold 2 is a triumph of Geometric ML and a paradigm shift in structural biology

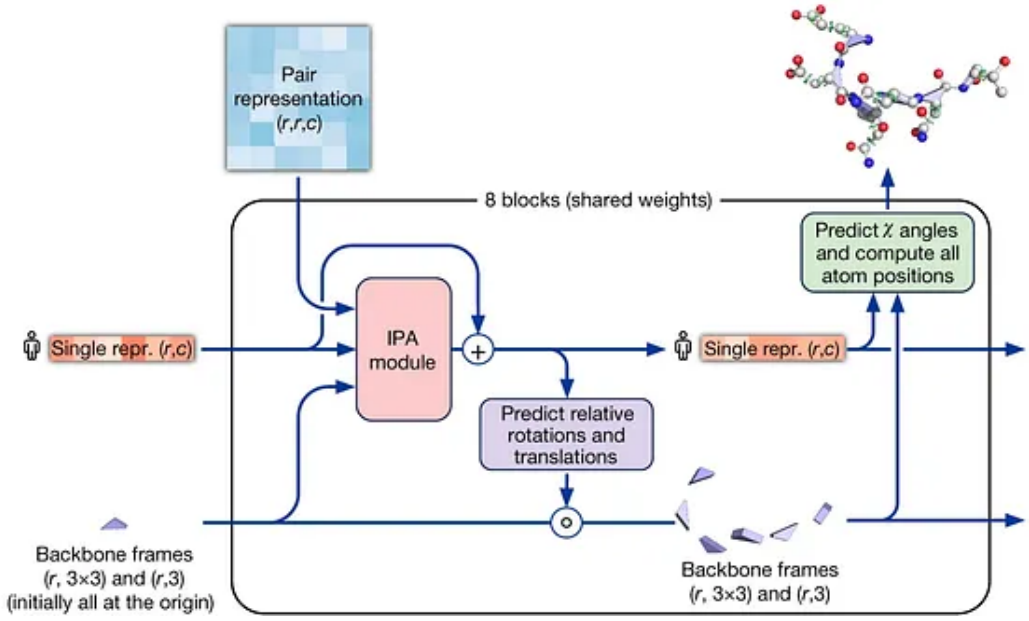
Just like Geometric Deep Learning might have been one of the defining trends of the past year (we take this opportunity to shamelessly advertise our own efforts in this direction culminating in the eponymous proto-book we wrote with Joan Bruna and Taco Cohen), there is little doubt about one particular groundbreaking instance of GDL in 2021. We are talking, of course, about AlphaFold 2 [95]: an architecture for protein structure prediction that dominated the Critical Assessment of protein Structure Prediction (CASP14) competition, a kind of "ImageNet moment" for structural biologists. At the heart of AlphaFold 2's modules is *invariant point attention*, which explicitly takes into account the proteins' geometry. We caught up with one of the joint first authors of AlphaFold 2, for a first-hand account of this achievement and follow-up work.

> "Deep Learning has been a core tool in protein structure prediction over the last few years." — Alexander Pritzel, Research Scientist, DeepMind.

"In 2021, we have released our model AlphaFold 2 [95] along with source code and a database of predictions for a large number of model organisms which we plan to extend to a large fraction of known proteins. This has since accelerated progress in experimental protein structure determination by allowing researchers to use predicted structures to interpret their data [96]. There have also been initial steps towards mapping out interaction networks between proteins, for example applying AlphaFold 2 to understand eukaryotic interactions as well as human protein interactions [97–98]. In 2022, I expect that the availability of very large datasets of very accurate protein structure predictions will spur new research in using protein structures to predict the functional properties of proteins.

Another area where I hope to see progress is in graph inference. For many real-world problems, the underlying graph is not known *a priori* or the natural graph like the molecular graph is not the right graph for message passing, so the graph is part of the prediction. The idea of inferring a graph by using a vague notion of transitivity is what motivated one of the core architectural components of AlphaFold.

In summary, in the last year, we have seen ML become a core tool in biology and I expect to see much more work in this space, hopefully paving the way for an atomistic understanding of cellular biology by employing a combination of machine learning and experimental methods. I also hope to see ML become a core tool in many other areas of scientific research."



AlphaFold 2 brought a breakthrough in predicting protein folding structure. At the core of this architecture is the "invariant point attention" module implementing geometric equivariance. Image from Jumper et al. [95].

The success of AlphaFold 2 caused many practitioners — and indeed, many of our interviewees — to reflect and re-evaluate some of their assumptions on geometric deep learning. Aasa Feragen follows up with a somewhat philosophical view of the implications of this breakthrough to differential geometry-inspired methods:

"2021 has seen a number of examples of how differential geometric problems have been provided with successful solutions. A high-profile example is AlphaFold 2 [95] with its invariant point attention. In other cases, differential geometry is not explicitly used: e.g. topology-aware loss functions [99] and registration models [100] that punish failure to preserve topology without actually quantifying it; or learning geometric invariances without the explicit use of geometry [101].

I find this an interesting lesson: while differential geometry may be fundamental and useful for understanding and phrasing a problem, effective and widely accessible solutions are frequently found without it. Is this a sign that complex mathematical models play a bigger role in approaching a problem than in finally

solving it? Or is it a sign that we are missing out on even better models by allowing a persistent divide between the practical and theoretical communities?"

While AlphaFold 2 is perhaps one of the most discussed ML models of 2021, the authors of this blog post note that in 2021 there has also been an impressive effort in David Baker's lab called RosettaFold [102]. It is inspired by AlphaFold's ideas and results, and it is also a geometric deep learning architecture based on SE(3)-equivariance [103].

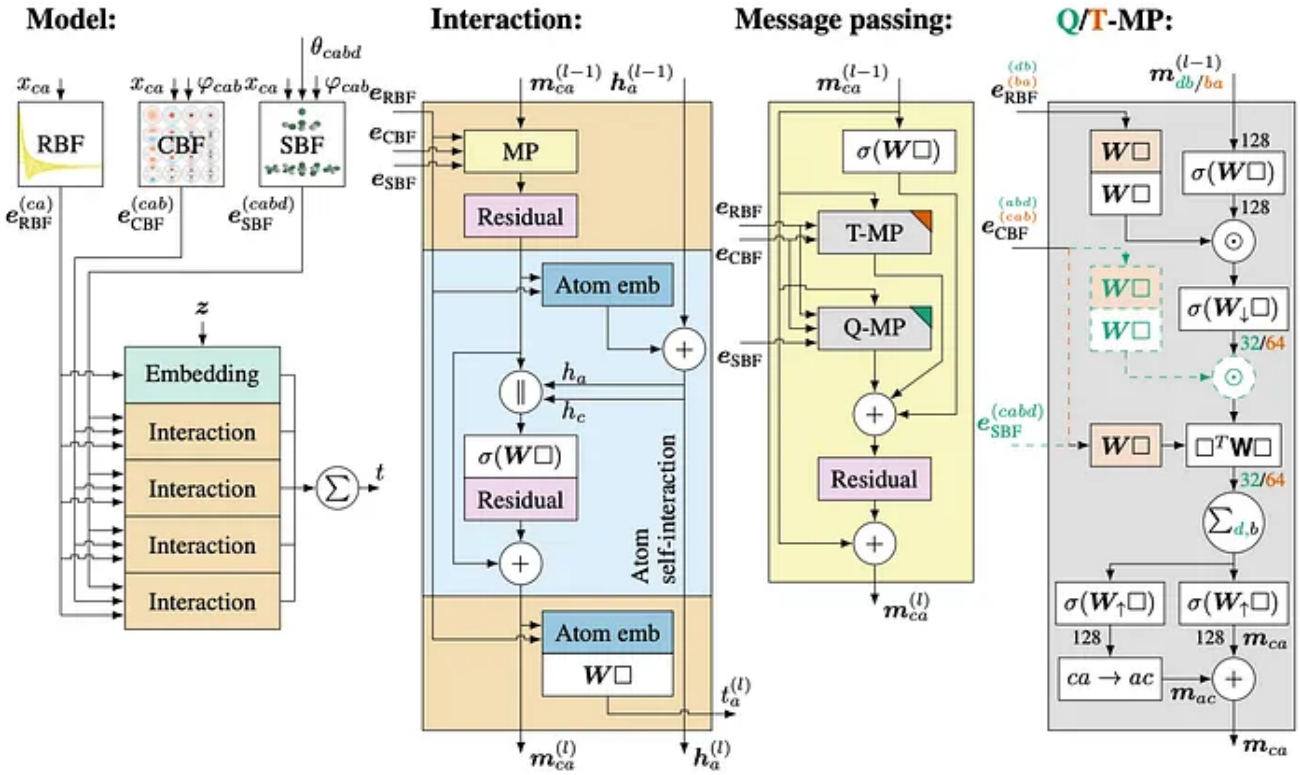
Drug discovery and design benefits from GNNs and their confluence with Transformers

The held-out predictions of AlphaFold 2 were so accurate that it was lauded as a "solution" to the protein folding problem. In this way, the very foundations of simulating macromolecules have been shaken. What can be said of smaller molecules, that are still the battle horses of drug therapies?

GNNs made some of their first notable impacts when applied on small molecules: inferring their solubility or toxicity [104], their quantum-chemical properties [105], or their potency as candidate drugs [106]. This is a wave that continued throughout 2021, as succinctly summarised by one of our contributors:

> "GraphML is continuously advancing the field of molecular modelling, with applications in both drug and material discovery." — Jian Tang, Assistant Professor, Mila

"In particular, I am very excited by the GemNet algorithm [107], which provides a general and expressive framework for 3D molecular structure modelling and achieves state-of-the-art performance in many tasks including molecular dynamics simulation and the open catalyst discovery competition. In 2022, I believe we will see more exciting progress along this direction, e.g., for protein modelling," Jian concludes.



The GemNet architecture is one of the highlights of 2021 in the buzzing field of molecular modelling with graph neural networks. Image from Klicpera et al. [107].

One of the key drivers behind research on GNNs for modelling small molecules is its immediate applicability for drug discovery pipelines. We contacted an expert at this specific intersection for comment, reaffirming this stance:

> "Drug discovery is one of the hottest topics for GNN research since molecules exhibit a graph structure and knowledge graphs are omnipresent in studying protein function and disease." — Dominique Beaini, Lead Deep Learning Researcher, Valence Discovery

"In drug discovery, one of the most challenging aspects of understanding chemistry is predicting quantum interactions and 3D structures from the molecular graph. Although there have been numerous works on the subject [108–109], the models are often too specific because they are designed solely to solve this problem, and it is unclear how well the 3D knowledge can be transferred to unrelated tasks [110]. Instead, a general GNN that can learn quantum mechanics will be able to transfer

learned embeddings to any new task, but this is only possible with a fully connected Transformer that can capture interactions beyond the molecular graph."

Indeed, 2021 has seen a flurry of concurrent research that successfully generalised Transformers to graphs. Dominique elaborates:

"At NeurIPS 2021, we have seen the first two Transformers that perform well on graph data, namely SAN [111] and Graphormer [112]. The latter showed that Transformers could better capture quantum mechanics by winning both the OGB LSC competition [29] and Facebook's catalyst competition [113], and that the learned embeddings could be transferred to unrelated biological tasks when they dominated the leaderboard for OGB-molPCBA [112]."

Modelling small molecules appears to be the area where graph Transformers have surfaced as a state-of-the-art approach — both theoretically and empirically. We reached out to one of the authors of the competition-winning Graphormer architecture for further insight:

> "In 2021, exciting progress has been made to adapt the Transformer architecture to the Graph ML domain [111–112,114–115] with different methods of incorporating graph structural information." — Tianle Cai, PhD student, Princeton University

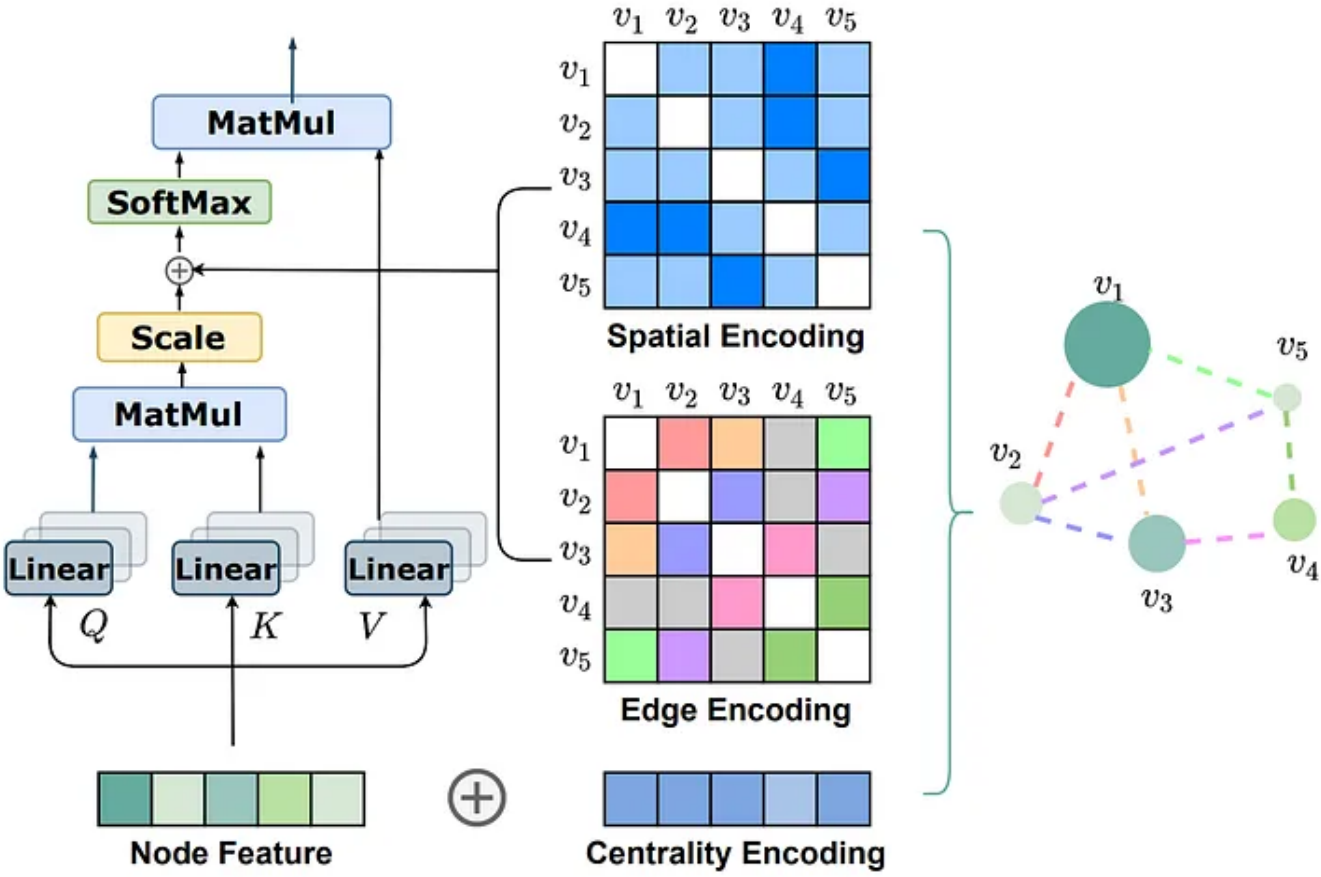
Tianle outlines several directions on which we can advance graph Transformers even further in 2022. One of them is GNN for solving scientific problems:

"The successes of Graphormer [112] on both quantum chemistry (KDD Cup — OGB-LSC) and molecule dynamics tasks (Open Catalyst Challenge) have shown the ability and potential of GNNs for scientific computing. Further, GNNs can inherently be designed [5] to enjoy several symmetry properties, which would make the model sound from a physics point of view.

The popularity of Transformer architecture in Natural Language Processing owes a great deal to the availability of models such as BERT, GPT, ViT, or CLIP pretrained on large datasets that appear to generalise well. The Graph Transformer architecture

has great expressive power, yet its generalisation ability should be guaranteed by sufficient amounts of data."

The computational complexity of full-graph attention is another issue: "Transformers have a quadratic time and space complexity w.r.t. the number of nodes, which makes it intractable in large graphs. Techniques for reducing this complexity should be developed to enable the use of Graph Transformers on larger graphs."



Boasting the familiar Transformer backbone, while packed with graph-specific innovation, Graphormer was one of the most empirically successful GNNs of 2021, especially on computational chemistry large-scale challenges. Figure from Ying et al. [112].

Tianle is excited to see further development of Graph Transformers, especially in the scientific domains: "I believe it is important to further scale up the current methods from both the data perspective (by pretraining on larger datasets or using the idea of Sim2Real to generate more data) and the model perspective (by designing more efficient architectures, such as linear Graph Transformers, or introducing graph coarsening or sampling techniques)."

Dominique affirms the expectation that more fundamental challenges surrounding Graph Transformers will be settled in 2022:

"Despite the successes of the early graph Transformers, they are still limited in expressive power, structural understanding of the graph, and scalability to larger graphs, but we expect to address these issues in 2022. Transformers will change how we can capture quantum and 3D information, how we generate molecules, and how we transfer knowledge across chemical tasks."

AI-first drug discovery is increasingly using Geometric and Graph ML

While the advancements in modelling small molecules have already yielded breakthroughs for the virtual drug screening pipeline [106], combining them with the structural biology results of AlphaFold can open up the doors to a plethora of important applications across the entire drug discovery pipeline. Many emerging startups in the drug discovery space are recognising this, and suggest that an "AI-first" approach to drug discovery is within reach. Nathan Benaich, a computational biologist turned venture capital investor at Air Street Capital, a firm focused on AI-first companies, elaborates on this from an entrepreneurial perspective:

"We've seen Geometric and Graph ML methods rip and replace classical methods on problems such as mesh-based simulations [38] and structure prediction for mRNAs [117] and proteins [95]. Going forward, I am anticipating Geometric and Graph ML to address a growing suite of problems across the natural and physical sciences in our efforts to decode their complexity. Tangibly, this could mean step changes to the effectiveness of AI-first drug and material design campaigns, particularly for multi-parameter optimisation, property prediction, and binding.

More fundamentally, I expect advances focused on overcoming the memory bottleneck of large Geometric and Graph ML as well as addressing the effect of depth on model performance versus the need for better architectures altogether."

This sentiment is mirrored by Dominique Beaini, who leads an applied research team at one such AI-first drug development startup:

"The recent success of AlphaFold to predict protein folding has fueled the community with hopes that machine learning will significantly impact biological research [95], and numerous companies such as Valence Discovery are building powerful drug discovery platforms supported by the latest advances in GNNs."

One of the most imminent ways to fuse these advances is to provide more accurate models of how molecules, small and large, interact with each other. So far, most emerging techniques focused on analysing molecules as if they are in a vacuum. Breaking away from this assumption could lead to some of the most exciting applications of graph representation learning to date, as highlighted by one of the most active GNN community members this previous year:

> "Graph ML has already transformed the analysis of individual proteins and especially small molecules. But what about modelling their interactions?" — Hannes Stärk, MSc student, TU Munich.

Hannes first detailedly recounts and amplifies many of the advances our other correspondents highlighted: "The power of Graph ML for molecular predictions is well established [26,117–118], and GNNs have achieved impressive successes in impactful domains such as antibiotics discovery [106]. Using the 3D geometry of molecules, Graph ML also advanced quantum chemistry through 3D GNNs such as GemNet [107] or molecular dynamics with NequIP [119]. E(3)-equivariant GNNs [5] and Geometric Deep Learning have become an important part of learning on molecules [120].

Similarly, 3D GNNs show strong results for protein representation learning with approaches to make their use computationally feasible for these large molecules [121]. Regarding Graph ML for proteins, I am, of course, also delighted about AlphaFold2 [95], which significantly advanced protein structure prediction and, therefore, the whole field of molecular biology."

From these foundations, Hannes proceeds to outline the excitement surrounding molecular interactions, and their implications for drug discovery: "Considering the now essential role of Graph ML for predictions about individual molecules, I think breakthroughs are on the horizon for modelling molecular interactions. While predictions about a molecule's inherent properties are interesting, I would imagine that there often is much more real-world significance to knowing, e.g., how proteins interact with others in our body or how a molecule interacts with a virus to inhibit its function. The research direction already started taking shape in 2021 with geometric deep learning approaches to

predict the atom configurations in which proteins interact with small molecules [122] or other proteins [14].

I am excited about the potential of Graph ML for, e.g., geometry-aware binding prediction or for projecting how the atoms of multiple molecules will move and how their conformations will change when they interact. These applications where more than a single molecule is modelled bring additional difficulties for including 3D information in a principled manner. As such, there are many open avenues for creative ideas to capture symmetries and prior geometric knowledge about specific molecular interactions. I think 2022 is perfectly set up to discover some of them, and with the field's massive potential for positive impacts, I cannot wait to see how Graph ML advances molecular interaction predictions."

Quantum ML benefits from graph-based methods

We finish with a field that is still a somewhat exotic niche to most ML people, but rapidly becoming less so, and entire industries are betting on seeing more of it a question of "when" rather than "if". This is, of course, quantum computing, and its more recent applications to machine learning. What does it all have to do with graphs? Guillaume Verdon, working at the intersection of these two fields, explains:

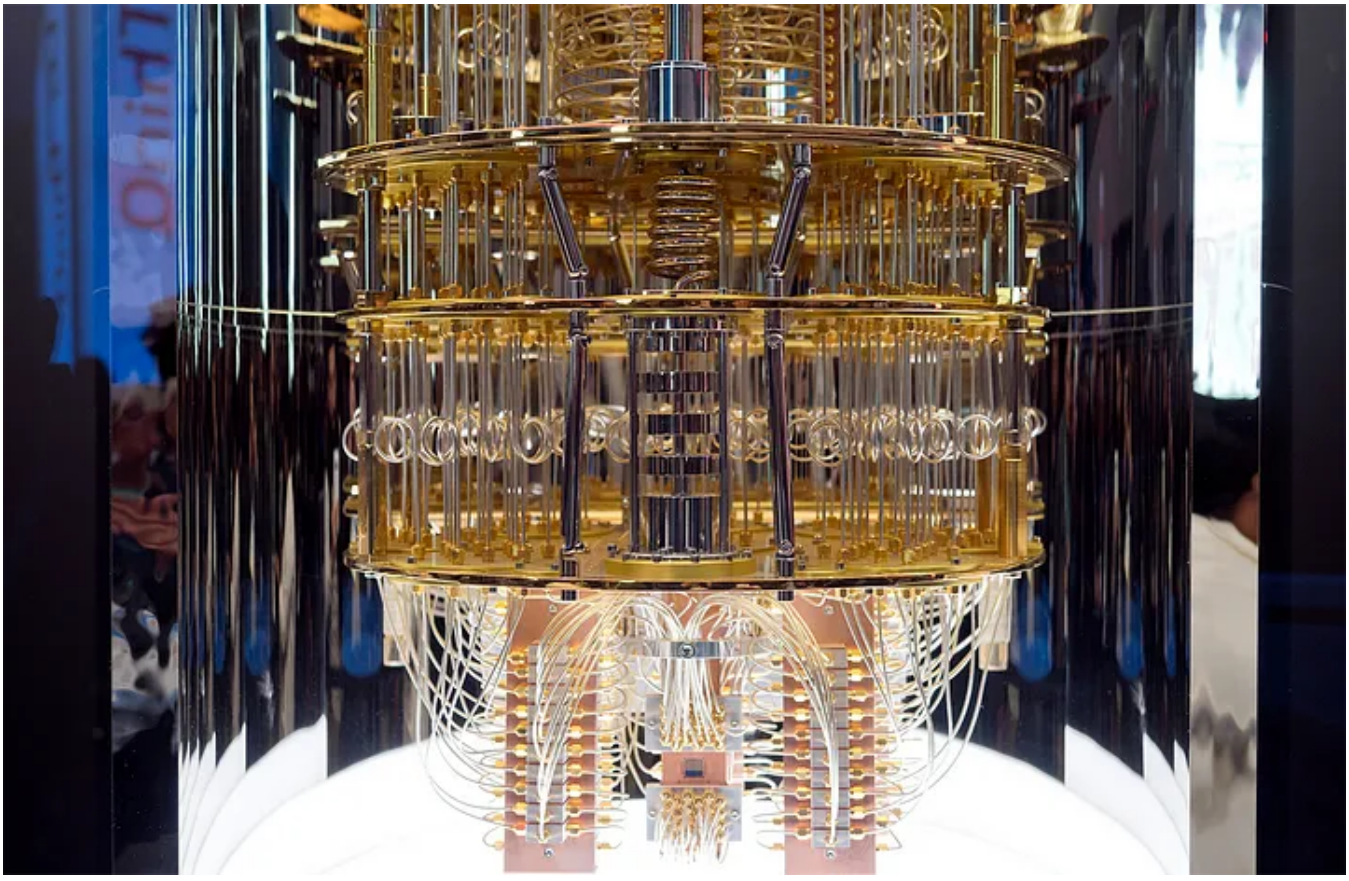
"Quantum Machine Learning (QML) [123], a nascent field of research at the intersection of ML and Quantum Computing, has seen a surge of interest both from academia and industry in recent years, especially with the advent of classically unsimulatable quantum computers [124]. One of the main attractions of QML is its ability to learn quantum models and representations from quantum data. As nature itself is quantum mechanical, learning quantum representations of data can help us extend the reach of classical ML to model nature at its fundamental level. In 2021, Google experimentally demonstrated an exponential scaling advantage for Quantum ML of quantum data on the Sycamore processor [125] and was a promising milestone for the field."

Beyond this early demonstration, several challenges for QML remain, a major hurdle being the design of quantum neural network architectures that are scalable, trainable, and provide strong generalization [126]. Quantum neural architectures which take inspiration from classical GNN's have recently popped up [127] to address this problem and have been successfully implemented in quantum hardware this past year [128].

A quantum system's evolution is typically governed by a so-called Hamiltonian, which for most systems has a locality of coupling between sub-systems that corresponds to a graph. As such, adding graph structural inductive biases to our quantum architectures is very natural for modelling both dynamics [127] and equilibrium properties [129–130] of quantum systems."

And so we wrap up on an optimistic and ambitious note:

> "An exciting direction to extend this graph quantum ML research in 2022 would be to explore beyond graphs and towards a quantum version of Geometric Deep Learning theory, as quantum physical systems often possess rich and esoteric group symmetries which could be leveraged for quantum architecture design, thereby furthering our ability to generatively model such systems using quantum computers." — Guillaume Verdon, Quantum ML Lead, Alphabet X



Will quantum computers be the next frontier for Geometric ML? Image: Shutterstock.

[1] M. Boguñá et al., Network geometry (2021) Nature Reviews Physics 3:114–135.

[2] Q. Liu, M. Nickel, D. Kiela, Hyperbolic Graph Neural Networks (2019) NeurIPS.

[3] M. Law. Ultrahyperbolic Neural Networks (2021) NeurIPS.

[4] Y. Zhang et al., Lorentzian Graph Convolutional Networks (2021) WWW.

[5] V. G. Satorras, E. Hoogeboom, M. Welling, E(n) equivariant graph neural networks (2021) ICML.

[6] J. Topping et al., Understanding over-squashing and bottlenecks on graphs via curvature (2022) ICLR. See also an accompanying blog post.

[7] H. Maron et al., Invariant and equivariant graph networks (2019) ICLR.

[8] T. Cohen et al., Gauge Equivariant Convolutional Networks and the Icosahedral CNN (2019) ICML.

[9] P. de Haan, T. Cohen, and M. Welling, Natural Graph Networks (2020) NeurIPS.

[10] E. H. Thiede, W. Zhou, and R. Kondor, Autobahn: Automorphism-based Graph Neural Nets (2021) NeurIPS.

[11] E. Moshe et al., PDE-GCN: Novel architectures for graph neural networks motivated by partial differential equations (2021) NeurIPS.

[12] B. P. Chamberlain et al., Beltrami flow and neural diffusion on graphs (2021), NeurIPS.

[13] U. Alon and E. Yahav, On the bottleneck of graph neural networks and its practical implications (2021) ICLR.

[14] O.-E. Ganea et al., Independent SE(3)-equivariant models for end-to-end rigid protein docking (2022) ICLR.

[15] N. S. Detlefsen, S. Hauberg, and W. Boomsma, What is a meaningful representation of protein sequences? (2021) arXiv:2012.02679.

[16] R. Sengers, L. Florack, and A. Fuster. Geodesic tubes for uncertainty quantification in Diffusion MRI (2021) IPMI.

[17] P. J. Kelly, On isometric transformations (1942) PhD Dissertation.

[18] B. Chamberlain et al., GRAND: Graph Neural Diffusion (2021) ICML. See also an accompanying blog post.

[19] D. I. Shuman et al., The emerging field of Signal Processing on Graphs (2013), IEEE Signal Processing Magazine 30(3):83–98.

[20] M. Weiler et al., Coordinate independent convolutional networks — Isometry and gauge equivariant convolutions on Riemannian manifolds (2021), arXiv:2106.06020.

[21] D. Kreutzer et al., Rethinking Graph Transformers with spectral attention (2021), NeurIPS.

[22] M. M. Bronstein et al., Geometric deep learning: going beyond Euclidean data (2017), IEEE Signal Processing Magazine 34(4):18–42.

[23] J. Bruna et al., Spectral Networks and Locally Connected Networks on Graphs (2014) ICLR.

[24] D. C. McNamee et al., Flexible modulation of sequence generation in the entorhinal–hippocampal system (2021) Nature Neuroscience 24:851–862.

[25] P. Piray and N. D. Daw, Linear reinforcement learning in planning, grid fields, and cognitive control (2021) Nature Communications 12.

[26] K. L. Stachenfeld, M. M. Botvinick, and S. J. Gershman, The hippocampus as a predictive map (2017) Nature Neuroscience 20:1643–1653.

[27] A. Derrow-Pinon et al., ETA Prediction with Graph Neural Networks in Google Maps (2021) arXiv:2108.11482. See also accompanying blog post.

[28] T. Pfaff et al., Learning mesh-based simulation with Graph Networks (2020) NeurIPS.

[29] W. Hu et al., OGB-LSC: A Large-Scale Challenge for Machine Learning on Graphs (2021) arXiv:2103.09430.

[30] J. Wang et al., scGNN is a novel graph neural network framework for single-cell RNA-Seq analyses (2021) Nature Communications 12.

[31] A. Bessadok, M. A. Mahjoub, and I. Rekik, Brain multigraph prediction using topology-aware adversarial Graph Neural Network (2021) arXiv:2105.02565.

[32] X. Li et al., BrainGNN: Interpretable Brain Graph Neural Network for fMRI analysis (2021), Medical Image Analysis 74.

[33] L. Torres et al., The Why, How, and When of representations for complex systems (2021) SIAM Review 63(3):435–485.

[34] F. Battiston et al., Networks beyond pairwise interactions: Structure and dynamics (2020) Physics Reports 874:1–92.

[35] A. R. Benson, D. F. Gleich, and D. J. Higham: Higher-order network analysis takes off, fueled by classical ideas and new data (2021) SIAM News (see also extended version).

[36] T. Eliassi-Rad et al., Higher-order graph models: From theoretical foundations to machine learning (2021) Report from Dagstuhl Seminar 21352.

[37] C. Bodnar, F. Frasca, et al., Weisfeiler and Lehman go topological: Message Passing Simplicial Networks (2021) ICML.

[38] C. Bodnar, F. Frasca, et al., Weisfeiler and Lehman go cellular: CW Networks (2021) NeurIPS.

[39] S. Ebli, M. Defferrard, and G. Spreemann, Simplicial Neural Networks (2020) NeurIPS Workshop on Topological Data Analysis and Beyond.

[40] E. Bunch et al., Simplicial 2-Complex Convolutional Neural Nets (2020) NeurIPS Workshop on Topological Data Analysis and Beyond.

[41] T. M. Roddenberry, N. Glaze, and Santiago Segarra, Principled Simplicial Neural Networks for trajectory prediction (2021) ICML.

[42] M. Hajij, K. Istvan, and G. Zamzmi, Cell Complex Neural Networks (2020) NeurIPS Workshop on Topological Data Analysis and Beyond.

[43] M. Hajij, G. Zamzmi, and X. Cai, Simplicial Complex Representation Learning (2021) arXiv:2103.04046.

[44] A. D. Keros, V. Nanda, and K. Subr, Dist2Cycle: A Simplicial Neural Network for homology localization (2021) AAAI.

[45] Y. Chen, Y. R. Gel, and H. V. Poor, BScNets: Block Simplicial Complex Neural Networks (2022) AAAI.

[46] J. G. Lambourne et al., BRepNet: A topological message passing system for solid models (2021) CVPR.

[47] J. M. Curry, Sheaves, Cosheaves and Applications (2014) PhD Dissertation.

[48] J. Hansen and T. Gebhart, Sheaf Neural Networks (2020) NeurIPS Workshop on Topological Data Analysis and Beyond.

[49] A. Parada-Mayorga et al., Quiver Signal Processing (2020) arXiv:2010.11525.

[50] D. Wu et al., Quantifying uncertainty in deep spatiotemporal forecasting (2021) KDD.

[51] C. Shang, J. Chen, and J. Bi, Discrete graph structure learning for forecasting multiple time series (2020) ICLR.

[52] R. Walters, J. Li, R. Yu, Trajectory prediction using equivariant continuous convolution (2021) ICLR.

[53] N. Dehmami et al., Automatic symmetry discovery with Lie Algebra Convolutional Network (2021), NeurIPS.

[54] A. D'Amour et al., Underspecification Presents Challenges for Credibility in Modern Machine Learning (2020) arXiv:2011.03395.

[55] W. Hu et al., Open Graph Benchmark: Datasets for Machine Learning on Graphs (2021) NeurIPS.

[56] B. Bevilacqua, Y. Zhou, and B. Ribeiro, Size-Invariant Graph Representations for Graph Classification Extrapolations (2021) ICML.

[57] S. C. Mouli and B. Ribeiro, Neural Networks for Learning Counterfactual G-Invariances from Single Environments (2021) ICLR.

[58] V. Veitch et al., Counterfactual Invariance to Spurious Correlations in Text Classification (2021) NeurIPS.

[59] G. Yehudai et al., From Local Structures to Size Generalization in Graph Neural Networks (2021) ICML.

[60] L. Lovasz and B. Szegedy, Limits of dense graph sequences (2006) Journal of Combinatorial Theory, Series B.

[61] Anonymous, Discovering Invariant Rationales for Graph Neural Networks (2021), https://openreview.net/forum?id=hGXij5rfiHw

[62] L. Cotta, C. Morris, and B. Ribeiro, Reconstruction for Powerful Graph Representations (2021) NeurIPS.

[63] B. Bevilacqua et al., Equivariant Subgraph Aggregation Networks (2022) ICLR.

[64] P. A. Papp et al., DropGNN: Random dropouts increase the expressiveness of Graph Neural Networks (2021) arXiv:2111.06283.

[65] G. Bouritsas et al., Improving graph neural network expressivity via subgraph isomorphism counting (2020) arXiv:2006.09252; see an accompanying post.

[66] L. Zhao et al., From stars to subgraphs: Uplifting any GNN with local structure awareness (2021) arXiv:2110.03753.

[67] B. Srinivasan and B. Ribeiro, On the equivalence between positional node embeddings and structural graph representations (2020) ICLR 2020.

[68] M. Zhang et al., Labeling Trick: A theory of using Graph Neural Networks for multi-node representation learning (2021) NeurIPS.

[69] V. P. Dwivedi et al., Graph Neural Networks with learnable structural and positional representations (2021) arXiv:2110.07875.

[70] T. Joachims et al., Recommendations as treatments (2021) AI Magazine 42(3):19–30.

[71] Z. Zhu et al., Neural Bellman-Ford Networks: A general graph neural network framework for link prediction (2021) NeurIPS.

[72] A. Tamar et al., Value iteration networks (2016) NeurIPS.

[73] G. Farquhar et al., TreeQN and ATreeC: Differentiable tree-structured models for deep reinforcement learning (2018) ICLR.

[74] T. Wang et al., Nervenet: Learning structured policy with graph neural networks (2018) ICLR.

[75]: V. Kurin et al., My body is a cage: the role of morphology in graph-based incompatible control (2021) ICLR.

[76] C. Blake et al., Snowflake: Scaling GNNs to High-Dimensional Continuous Control via Parameter Freezing (2021) NeurIPS.

[77] W. Shang et al., Agent-Centric Representations for Multi-Agent Reinforcement Learning (2021) arXiv:2104.09402.

[78] S. Li et al., Deep implicit coordination Graphs for Multi-agent Reinforcement Learning (2021) AAMAS.

[79] K. S. Luck, R. Calandra, and M. Mistry, What robot do I need? Fast co-adaptation of morphology and control using Graph Neural Networks (2021) arXiv:2111.02371.

[80] D. J. Hejna III, P. Abbeel, and L. Pinto: Task-agnostic morphology evolution (2021) ICLR.

[81] D. Yoon et al., Winning the L2RPN Challenge: Power grid management via semi-Markov afterstate actor-critic (2020) ICLR.

[82] Q. Cappart et al., Combinatorial optimization and reasoning with graph neural networks (2021) arXiv:2102.09544.

[83] A. Z. Wagner, Constructions in combinatorics via neural networks (2021) arXiv:2104.14516.

[84] A. I. Deac et al., Neural algorithmic reasoners are implicit planners (2021) NeurIPS.

[85] Z. Jiang et al., Grid-to-Graph: Flexible spatial relational inductive biases for reinforcement learning (2021) AAMAS.

[86] H. Zhu et al., Network planning with deep reinforcement learning (2021) SIGCOMM.

[87] C. Cummins et al., CompilerGym: Robust, performant compiler optimization environments for AI research (2021) arXiv:2109.08267.

[88] A. Vaswani et al., Attention is all you need (2017) NIPS.

[89] L. Chen et al., Decision Transformer: Reinforcement learning via sequence modeling (2021) arXiv:2106.01345.

[90] M. Janner, Q. Li, and S. Levine, Offline reinforcement learning as one big sequence modeling problem (2021) NeurIPS.

[91] P. Veličković et al., Graph attention networks (2018) ICLR.

[92] S. Hu et al., UPDeT: Universal multi-agent RL via policy decoupling with Transformers (2021) ICLR.

[93] Y. Niu, R. Paleja, and M. Gombolay, Multi-agent graph-attention communication and teaming (2021) AAMAS.

[94] E. van der Pol et al., Multi-agent MDP homomorphic networks (2022) ICLR.

[95] J. Jumper et al., Highly accurate protein structure prediction with AlphaFold, Nature 596:583–589, 2021.

[96] G. Masrati et al., Integrative structural biology in the era of accurate structure prediction (2021) Journal of Molecular Biology 433(20).

[97] I. R. Humpreys et al., Computed structures of core eukaryotic protein complexes (2021) Science 374(6573).

[98] D. F. Burke et al., Towards a structurally resolved human protein interaction network (2021) bioarXiv:2021.11.08.467664.

[99] S. Shit et al. clDice-a Novel Topology-Preserving Loss Function for Tubular Structure Segmentation (2021) CVPR.

[100] P. S. Czolbe, A. Feragen, and O. Krause, Spot the difference: Detection of topological changes via geometric alignment (2021) NeurIPS.

[101] P. E. Schwöbel et al., Last layer marginal likelihood for invariance learning (2021) arXiv:2106.07512.

[102] M. Baek et al., Accurate prediction of protein structures and interactions using a three-track neural network, Science 373:871–876, 2021.

[103] F. B. Fuchs et al., SE(3)-Transformers: 3D roto-translation equivariant attention networks (2020) NeurIPS.

[104] D. Duvenaud et al., Convolutional networks on graphs for learning molecular fingerprints (2015) NeurIPS.

[105] J. Gilmer et al., Neural message passing for quantum chemistry (2017) ICML

[106] J. M. Stokes et al., A deep learning approach to antibiotic discovery (2020) Cell 180(4):688–702.

[107] J. Klicpera, F. Becker, and S. Günnemann, GemNet: Universal directional graph neural networks for molecules (2021) arXiv:2106.08903.

[108] S. Luo et al., Predicting molecular conformation via dynamic graph score matching (2021) NeurIPS.

[109] C. Shi C et al., Learning gradient fields for molecular conformation generation (2021) ICLR.

[110] H. Stärk et al., 3D Infomax improves GNNs for molecular property prediction (2021) arXiv:2110.04126.

[111] D. Kreuzer D et al., Rethinking Graph Transformers with spectral attention (2021) NeurIPS.

[112] C. Ying C et al., Do Transformers really perform bad for graph representation? (2021) NeurIPS.

[113] L. Chanussot et al., The Open Catalyst 2020 (OC20) Dataset and Community Challenges (2021) ACS Catal 11:6059–6072.

[114] V. P. Dwivedi and X. Bresson. A generalization of transformer networks to graphs (2021) Workshop on Deep Learning on Graphs: Methods and Applications.

[115] G. Mialon et al., Graphit: Encoding graph structure in transformers (2021) arXiv:2106.05667.

[116] T. Pfaff et al., Learning mesh-based simulation with graph networks (2021) ICLR.

[117] J. Pan, Deep learning boosts RNA prediction (2021) Nature Computational Science 1(564).

[117] K. Yang et al., Analyzing learned molecular representations for property prediction (2019) J. Chem. Inf. Model 59(8):3370–3388.

[118] O.-E. Ganea et al., GeoMol: Torsional geometric generation of molecular 3D conformer ensembles (2021) NeurIPS.

[119] S. Batzner et al., E(3)-equivariant Graph Neural Networks for data-efficient and accurate interatomic potentials (2021) arXiv:2101.03164.

[120] K. Atz et al., Geometric deep learning on molecular representations (2021) Nature Machine Intelligence 3:1023–1032.

[121] V. R. Somnath et al., Multi-scale representation learning on proteins (2021) NeurIPS.

[122] O. Méndez-Lucio et al., A Geometric Deep Learning approach to predict binding conformations of bioactive molecules (2021) Nature Machine Intelligence 3:1033–1039.

[123] M. Broughton et al., Tensorflow quantum: A software framework for quantum machine learning (202) arXiv:2003.02989.

[124] F. Arute et al., Quantum supremacy using a programmable superconducting processor (2019) Nature 574(7779): 505–510.

[125] H.-Y. Huang et al., Quantum advantage in learning from experiments (2021) arXiv:2112.00778.

[126] J. R. McClean et al., Barren plateaus in quantum neural network training landscapes (2018) Nature Communications 9:1–6.

[127] G. Verdon et al., Quantum graph neural networks (2019) arXiv:1909.12264.

[128] L.-P. Henry et al., Quantum evolution kernel: Machine learning on graphs with programmable arrays of qubits (2021) Physical Review A 104(3).

[129] G. Verdon et al., Quantum Hamiltonian-based models and the variational quantum thermalizer algorithm (2019) arXiv:1910.02071.

[130] C. Zoufal, A. Lucchi, and S. Woerner, Variational quantum Boltzmann machines (2021) Quantum Machine Intelligence 3(1):1–15.

This post was written together with Petar Veličković, based on contributions from Dominique Beaini (Valence Discovery), Nathan Benaich (Air Street Capital), Cristian Bodnar (University of Cambridge), Tianle Cai (Princeton University), Tina Eliassi-Rad (Northeastern University), Aasa Feragen (University of Copenhagen), Pim de Haan (University of Amsterdam), Francesco Di Giovanni (Twitter), Vitaly Kurin (University of Oxford), Haggai Maron (Nvidia), Alexander Pritzel (DeepMind), Bruno Ribeiro (Purdue University), Kim Stachenfeld (DeepMind), Hannes Stärk (TU Munich), Jian Tang (Mila / HEC Montréal), Elise van der Pol (University of Amsterdam), Pierre Vandergheynst (EPFL), Guillaume Verdon (Alphabet X), Rose Yu (UC San Diego), and Melanie Weber (University of Oxford). Needless to say, all the credit goes to the aforementioned people, whereas any criticism should be the sole responsibility of the authors. We are also grateful to Ben Chamberlain, Fabrizio Frasca, Maria Gorinova, and Emanuele Rossi for proofreading this post.

For additional articles about deep learning on graphs, see Michael's other posts in Towards Data Science, subscribe to his posts and YouTube channel, get Medium membership, or