Newsjournal of the Society for Industrial and Applied Mathematics sinews.siam.org

siam news

Volume 55/ Issue 4
May 2022

# Protecting Privacy with Synthetic Data

*By Matthew R. Francis*

Researchers across every scientific discipline need complete and reliable data sets to draw trustworthy conclusions. However, publishing all data from a given study can be undesirable. For example, medical data in particular include personal information that—if published in full—would violate patients' privacy and potentially expose them to harm. Similarly, many studies in the social sciences include demographic or geographical data that could easily be exploited by unscrupulous parties.

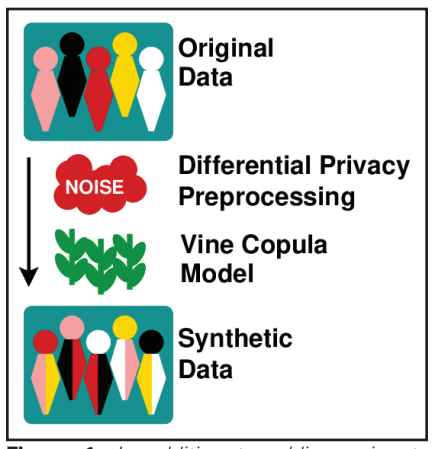



**Figure 1.** *In addition to adding noise to the data set, Sébastien Gambs' differential privacy-based method processes it with an information theory algorithm to obtain synthetic data that—in principle—shields the privacy of the people involved. Figure courtesy of the author.*

In short, researchers must strike a delicate balance between publishing enough data to verify their conclusions and protecting the privacy of the people involved. Unfortunately, multiple studies have shown that simply anonymizing the data—by removing individuals' names before publication, for instance—is insufficient, as outsiders can use context clues to reconstruct missing information and expose research subjects. "We want to generate synthetic data for public release to replace the original data set," Bei Jiang of the University of Alberta said. "When we design our framework, we have this main goal in mind: we want to produce the same inference results as in the original data set."

In contrast with falsified data, which is one of the deadliest scientific sins, researchers can generate *synthetic* data directly from original data sets. If the construction process is done properly, other scientists can then analyze this synthetic data and trust that their conclusions are no different from what they would have obtained with full access to the original raw data — ideally, at least. "When you [create] synthetic data, what does it mean to be private yet realistic?" Sébastien Gambs of the University of Québec in Montréal asked. "It's still an open research question."

During the 2022 American Association for the Advancement of Science Annual Meeting,[1] which took place virtually in February, Jiang and Gambs each presented formal methods for the generation of synthetic data that ensure privacy. Their models draw from multiple fields to address challenges in the era of big data, where the stakes are higher than ever. "There is always a trade-off between utility and risk," Jiang said. "If you want to protect people [who] are at a higher risk, then you perturb their data. But the utility will be lowered the more you perturb. A better approach is to account for their risks to begin with."

Unfortunately, malicious actors have access to the same algorithmic tools as researchers. Therefore, protection of confidentially also involves testing synthetic data against the types of attacks that such players might utilize. "In practice, this helps one really understand the translation between an abstract privacy parameter and a practical guarantee," Gambs said. In other words, the robustness of a formal mathematical model is irrelevant if the model is not well implemented.

[1] https://aaas.confex.com/aaas/2022/meetingapp.cgi

## Differential Privacy Made Simple(r)

Gambs and his collaborators turned to differential privacy: a powerful mathematical formalism that in principle is the best available technique for securing confidentiality. However, the approach is also complex and difficult to implement without a high degree of statistical knowledge. To smooth the

*See* **Synthetic Data** *on page 3*

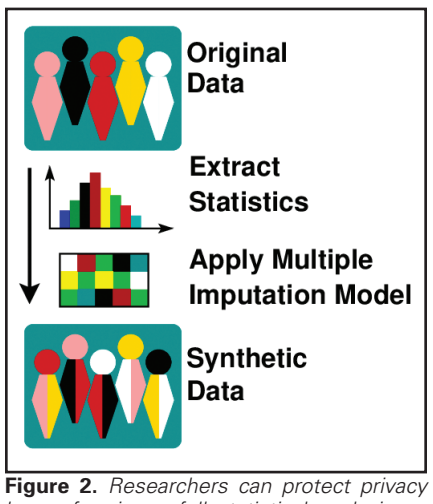



**Figure 2.** *Researchers can protect privacy by performing a full statistical analysis on the original data set, then using a missing-data algorithm called multiple imputation to construct a synthetic data set that has exactly the same statistical characteristics. Figure courtesy of the author.*



# Proving Existence Is Not Enough: Mathematical Paradoxes Unravel the Limits of Neural Networks in Artificial Intelligence

*By Vegard Antun, Matthew J. Colbrook, and Anders C. Hansen*

The impact of deep learning (DL), neural networks (NNs), and artificial intelligence (AI) over the last decade has been profound. Advances in computer vision and natural language processing have yielded smart speakers in our homes, driving assistance in our cars, and automated diagnoses in medicine. AI has also rapidly entered scientific computing. However, overwhelming amounts of empirical evidence [3, 8] suggest that modern AI is often non-robust (unstable), may generate hallucinations, and can produce nonsensical output with high levels of prediction confidence (see Figure 1). These issues present a serious concern for AI use within legal frameworks. As stated by the European Commission's Joint Research Centre, *"In the light of the recent advances in AI, the serious negative consequences of its use for EU citizens and organisations have led to multiple initiatives [...] Among the identified requirements, the concepts of robustness and explainability of AI systems have emerged as key elements for a future regulation."*[1]

Robustness and trust of algorithms lie at the heart of numerical analysis [9]. The *lack* of robustness and trust in AI is hence the Achilles' heel of DL and has become a serious political issue. Classical approximation theorems show that a continuous function can be approximated arbitrarily well by a NN [5]. Therefore, stable problems that are described by stable functions can be solved stably with a NN. These results inspire the following fundamental question: *Why does DL lead to unstable methods and AI-generated hallucinations, even in scenarios where we can prove that stable and accurate NNs exist?*

[1] https://publications.jrc.ec.europa.eu/repository/handle/JRC119336

Our main result reveals a serious issue for certain problems; while stable and accurate NNs may provably exist, no training algorithm can obtain them (see Figure 2, on page 4). As such, existence theorems on approximation qualities of NNs (e.g., universal approximation) represent only the first step towards a complete understanding of modern AI. Sometimes they even provide overly optimistic estimates of possible NN achievements.

## The Limits of AI: Smale's 18th Problem

The strong optimism that surrounds AI is evident in computer scientist Geoffrey Hinton's 2017 quote: *"They should stop training radiologists now."*[2] Such optimism is comparable to the confidence that surrounded mathematics in the early 20th century, as summed up in David Hilbert's sentiment: *"Wir müssen wissen. Wir werden wissen"* ["We must know. We will know"].

Hilbert believed that mathematics could prove or disprove any statement, and that there were no restrictions on which problems algorithms could solve. The seminal contributions of Kurt Gödel [7] and Alan Turing [12] turned Hilbert's idealism upside down by establishing paradoxes that expedited impossibility results about the feasible achievements of mathematics and digital computers.

A similar program on the boundaries of AI is necessary. Stephen Smale already suggested such a program in the 18th problem on his list of mathematical problems for the 21st century: *What are the limits of AI?* [11].

*See* **Mathematical Paradoxes** *on page 4*

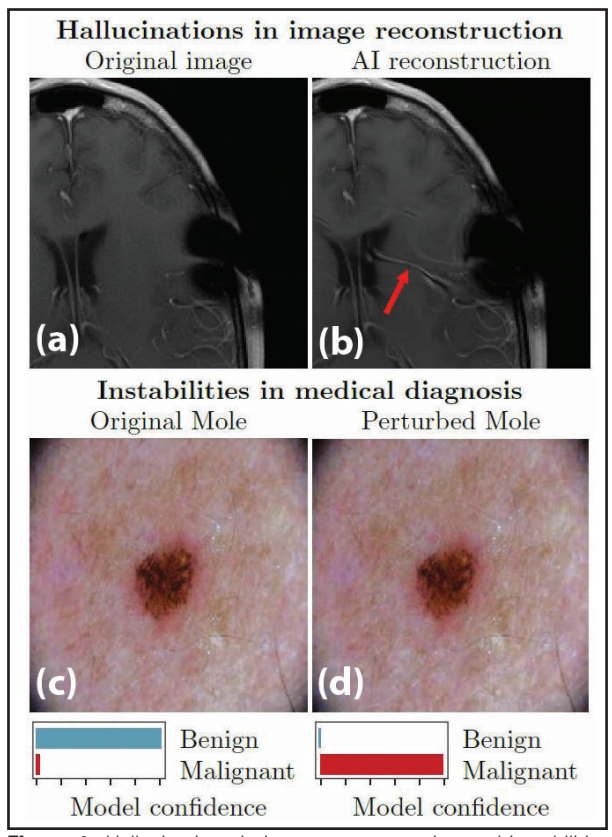


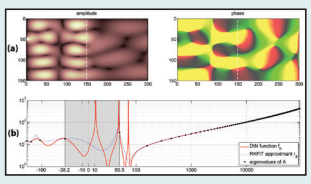

**Figure 1.** *Hallucinations in image reconstruction and instabilities in medical diagnoses.* **1a.** *The correct, original image from the 2020 fastMRI Challenge.* **1b.** *Reconstruction by an artificial intelligence (AI) method that produces an incorrect detail (AI-generated hallucination).* **1c.** *Dermatoscopic image of a benign melanocytic nevus, along with the diagnostic probability computed by a deep neural network (NN).* **1d.** *Combined image of the nevus with a slight perturbation and the diagnostic probability from the same deep NN. One diagnosis is clearly incorrect, but can an algorithm determine which one? Figures 1a and 1b are courtesy of the 2020 fastMRI Challenge [10], and 1c and 1d are courtesy of [6].*

[2] https://www.newyorker.com/magazine/2017/04/03/ai-versus-md

siam®
SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA

siam news
Volume 55/ Issue 4/ May 2022







# Building a Mathematical Toolbox for Biological Network Analysis

*By Lenore J. Cowen*

Large public databases that aggregate the many results of experiments in cell biology are increasingly expanding our knowledge of the relationships between human genes. Graph or network representations are appropriate for several types of data; for example, the classical protein-protein interaction (PPI) network places an edge between two genes if experimental evidence indicates that their associated proteins physically bind within the cell. Analysis of biological networks like PPI networks can yield powerful insights into important problems in functional genomics and predict previously unstudied genes that are related to key biological pathways that affect human health and disease.

Much like social networks, biological networks tend to be low-diameter, "small-world" networks. Homophily is a guiding organizational principle in most types of biological network data [10], meaning that genes typically associate with other genes that are involved with similar functional processes, pathways, and diseases. Some of the issues that researchers study in social and biological networks are directly analogous, and variants of diffusion processes—also called *network propagation* [6]—can immediately uncover important biological relationships and insights. However, several new problems require the development of novel domain-specific mathematical techniques in order to take full advantage of unique aspects of the biological network domain.

## Biological Networks Are Like Social Networks

The dense, low-diameter, small-world properties of both social and biological networks make it difficult for us to define a local neighborhood; notably, an ordinary shortest path distance is not sufficiently informative beyond direct neighbors. When we look to neighbors' neighbors (with a shortest path distance of two), we see an explosion in neighborhood size that can quickly reach a large fraction of the network in question for many types of association networks. We can resolve these ties in proximity to a more expressive and fine-grained measure of similarity via network propagation, which involves running short or lazy random walks from a gene of interest and observing the frequency at which different nodes are reached. A comprehensive survey of the multiple variants of network propagation that the computational biology community used to analyze PPI networks prior to 2019 is available in the literature [6].

Next we review one specific variant: diffusion state distance (DSD) [3, 5]. DSD is distinguished from other related measures in that it is a true distance metric that satisfies triangle inequality [3]. In particular, it is a graph embedding method that associates each node or gene with a vector. We define $H^k(u,v)$ as the expected number of times that the $k$-step symmetric random walk (starting with $k=0$) visits node $v$, and we classify the diffusion state vector that is associated with node $u$ as the vector of the $H^k(u,v_i)$ values for all the nodes $v_1,\ldots v_n$ in the graph. The DSD between nodes $a$ and $b$ is then

$$\mathrm{DSD}^k(a,b)=|H^k(a)-H^k(b)|_1,$$

where $|\cdot|_1$ represents the $L_1$ norm of the vectors. DSD satisfies triangle inequality [3]; these differences go to zero as the walk mixes for larger $k$, meaning that this measure converges as $k\to\infty$.

DSD also solves the ties in proximity problem. For a variety of different settings in the PPI network, we can show that the $t$ closest DSD neighbors define coherent functional neighborhoods in biological networks. This realization leads to improved performance on classical inference problems like functional label prediction [5]. For instance, a "double-spectral" DSD approach followed by spectral clustering yielded the best performance in the 2016 Disease Module Identification DREAM Challenge[1] [4], suggesting the presence of new genes that are involved in illnesses such as type 2 diabetes and Crohn's disease.

DSD is just one of numerous network propagation methods, and we could instead substitute alternative diffusion-based graph embedding techniques that have been previously proposed in the context of social networks. All of these diffusion-based method variants provide increased power and insight for the core classical problems that arise in biological networks, including functional label prediction, link prediction (i.e., predicting new PPIs that have not yet been experimentally observed), and disease module identification, which is directly analogous to the community detection problem in social networks (see Figure 1). The best such method is domain- and problem-specific. Customizing the appropriate embedding for the domain at hand then becomes an important and compelling challenge.

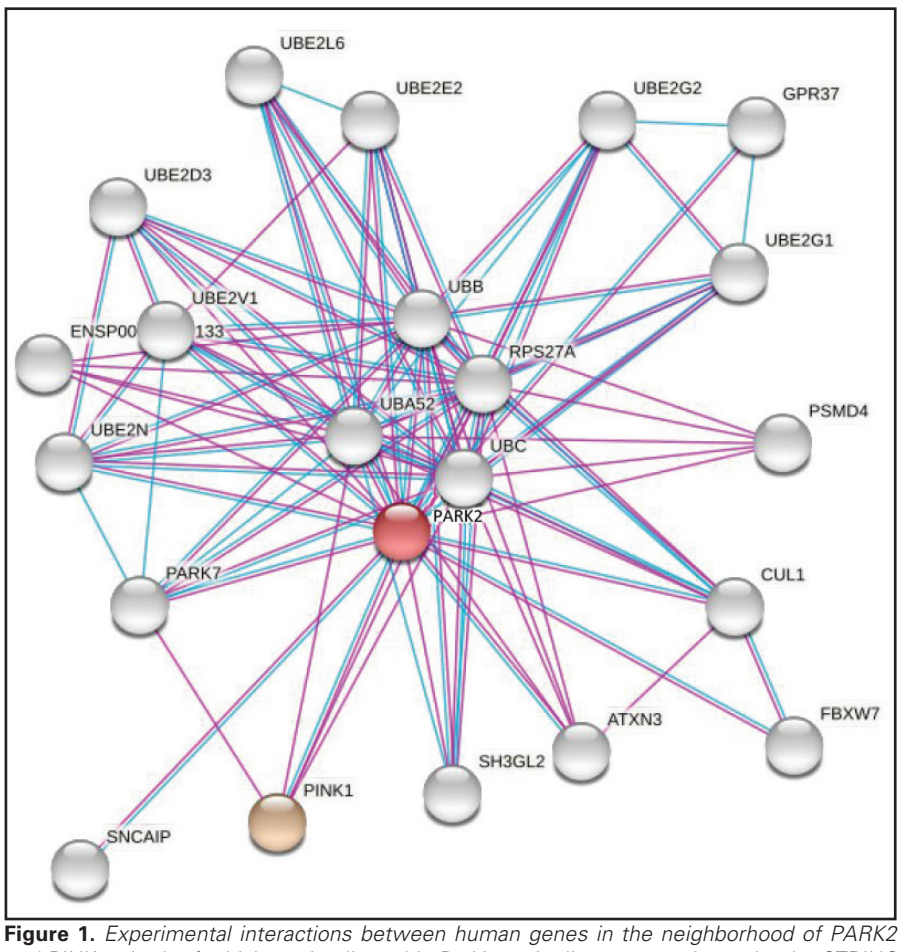



**Figure 1.** *Experimental interactions between human genes in the neighborhood of PARK2 and PINK1—both of which are implicated in Parkinson's disease—as shown by the STRING network [12]. This figure was generated from https://string-db.org with seed genes PARK2 (in red) and PINK1 (in light pink).*

## Biological Networks Differ From Social Networks

Although we can directly map many important problems in biological networks to well-studied social network analogs, three categories of domain-specific differences are worth highlighting: (i) The organizational principles of the networks themselves can be different, (ii) the nature and availability of ground-truth data to test methods may differ, and (iii) we can leverage the power

*See* **Biological Network** *on page 5*

[1] https://www.synapse.org/#!Synapse:syn6156761/wiki/400645

siam news

ISSN 1557–9573. Copyright 2022, all rights reserved, by the Society for Industrial and Applied Mathematics, SIAM, 3600 Market Street, 6th Floor, Philadelphia, PA 19104-2688; (215) 382-9800; siam@siam.org. To be published 10 times in 2021: January/February, March, April, May, June, July/August, September, October, November, and December. The material published herein is not endorsed by SIAM, nor is it intended to reflect SIAM's opinion. The editors reserve the right to select and edit all material submitted for publication.

**Advertisers:** For display advertising rates and information, contact Kristin O'Neill at marketing@siam.org.

**One-year subscription (nonmembers):** Electronic-only subscription is free. $73.00 subscription rate worldwide for print copies. SIAM members and subscribers should allow eight weeks for an address change to be effected. Change of address notice should include old and new addresses with zip codes. Please request address change only if it will last six months or more.

# A Firsthand Look at the Novel Ph.D. Program in Mathematical Modeling at Rochester Institute of Technology

*By Lina Sorg*

The innovative Ph.D. program in mathematical modeling[1] at Rochester Institute of Technology (RIT) accepted its initial cohort of students in 2017. This is the first and currently only Ph.D. track in the U.S. that focuses specifically on mathematical modeling, rather than applied mathematics or another variation thereof. Since its onset, the program—which seeks to fill a niche in the workforce for applied mathematicians and computational scientists—has grown steadily, exposing students to a wide variety of classes, research areas, and hands-on experiences. Applications doubled in the last year alone, likely because increasingly more students are encountering mathematical modeling at the undergraduate level.

*SIAM News* recently sat down with Nathan Cahill, who has served as the director of RIT's Ph.D. program in mathematical modeling since 2019. Nathan detailed the program's origins, explained the structure and organization of relevant coursework and research projects, discussed internship and career opportunities, and shared his vision for the program's continued evolution in the coming years.

***SIAM News*: Why did RIT decide to establish a Ph.D. program in mathematical modeling?**

**Nathan Cahill:** We owe a large part of our program's history to Sophia Maggelakis, current dean of RIT's College of Science and former head of the School of Mathematical Sciences.[2] She started her career as a mathematics faculty member and recognized a huge need in the workplace for the type of skills that are associated with a mathematical modeling education. Sophia realized that RIT could fill a gap by creating a program that specifically trains students to collaborate with experts in different fields and use mathematics to solve complicated, real-world problems. She broadened our Department of Mathematics and Statistics to a School of Mathematical Sciences, and brought in faculty who did not necessarily have traditional mathematics backgrounds but could involve students in research that applies math modeling to various settings. For instance, my Ph.D. is in engineering science.

In the early 2010s, we began mapping out a prospective Ph.D. program in mathematical modeling. It took a few years to put together a solid proposal and figure out how it would fit within RIT and the broader context of companies and government agencies that hire mathematicians. We finally got approval from the state of New York in 2016 and welcomed our first class of students in 2017.

**SN: Was it difficult to assemble a cohesive proposal since faculty backgrounds are so diverse?**

**NC:** We have a very heterogeneous group of faculty members in terms of research application areas, but we all draw on fundamental aspects of bringing mathematics to bear in real-world situations. We all tackle real-world problems in some capacity and work with experts who have domain-specific expertise that we don't necessarily possess. We think about how we can translate their problems into the language of mathematics; generate ideas based on mathematical tools; and propose mathematical methods to predict, analyze, and/or provide insights. Once we understood that we all have this common foundation, we saw the foundation as something that we can teach graduate students.

**SN: How has the program grown since its origin?**

**NC:** We started with about five students in 2017 and admitted five or six students for the first few years, but we'll probably admit eight or nine this fall. We tell students that it typically takes five years to get through the program; we're at that five-year point right now and our first two graduates have defended their dissertations. 28 Ph.D. students are currently enrolled across all levels, and 16 of them are women. We have a strong contingent of female faculty and want to build an accessible environment that promotes women in the pipeline of the workforce.

**SN: What is the program's structure?**

**NC:** The first year is entirely coursework. All students take one semester of numerical analysis and a two-semester sequence in mathematical modeling. We also require a one-credit graduate seminar course in both the fall and spring that teaches general research methods and exposes participants to different research areas. In addition, a high-performance computing course introduces students to distributed and parallel computing for various types of problems.

Candidates also choose a concentration in their first year, each of which has several core courses. Possible concentrations include biomedical mathematics, applied inverse problems and optimization, dynamical systems and fluid dynamics, and discrete mathematics, among other topics. Beyond these courses, students must enroll in three electives of their choosing; they can even take graduate courses in other departments.

During their second year, students must take at least three research credits, identify a faculty mentor, and carry out a research project. At the end of that year, an examination of this project determines whether they are ready to begin dissertation research. From then on, it's basically all dissertation research.

**SN: What makes RIT's math modeling Ph.D. different from a Ph.D. in applied mathematics?**

**NC:** There is of course some overlap, but the thing that sets us apart is that we want our students to learn the process of modeling itself and understand how to talk to experts in specific domain areas. We train students to work collaboratively across disciplines, tease out the problems that domain experts are trying to solve, and establish these problems in mathematical settings. So while our focus does involve applied mathematics, it's broader in the

*See* **Novel Ph.D. Program** *on page 6*

*Ph.D. student Jenna McDanold is a research assistant at Los Alamos National Laboratory (LANL), where she studies the way in which tree characteristics affect surface fuel buildup and wildfire behavior. LANL is currently supporting her as she completes her Ph.D. in mathematical modeling at Rochester Institute of Technology. Photo courtesy of LANL.*

[1] https://www.rit.edu/study/mathematical-modeling-phd

[2] https://www.rit.edu/science/school-mathematical-sciences

## Synthetic Data

*Continued from page 1*

edges, Gambs' team combined differential privacy with a probabilistic concept known as vine copula—drawn from a model that mathematician Abe Sklar first published in 1959 [4]—and applied the method to real data to demonstrate its usefulness.

In pure differential privacy applications,[2] researchers add a judicious amount of noise that is scaled by a parameter $\varepsilon$ to the original data set in order to disrupt correlations between variables that could identify sensitive information. Though this framework offers the strongest possible privacy guarantee in formal terms, it has a number of practical drawbacks. "One issue with differential privacy is keeping as much utility as possible," Gambs said. "It's abstract in the sense that the mathematics and the [privacy] guarantee are very formal. It's difficult for people who use this method to understand what it means to choose a particular value of $\varepsilon$ in terms of what they're protecting from privacy attacks."

Gambs and his colleagues wanted their toolkit to be extremely user-friendly and flexible so that researchers without expertise in formal privacy mathematics can apply it to their own data [1]. Their algorithm first preprocesses the raw data via differential privacy methods and a preselected noise "budget," then follows up with the vine copula technique on the scrambled data set to produce the final synthetic tables for analysis (see Figure 1, on page 1). The group also applies known types of attacks to their own synthetic data sets to ensure that their security levels are sufficient. "When you use differential privacy, you have some assumptions about the distribution of the data that are not necessarily true in real life," Gambs said. "Instead, you might have a subpopulation [with] a high correlation between profiles. If you just rely on the theoretical guarantee of differential privacy but don't do any practical privacy attacks, you might miss this kind of problem in the data."

### Striving for Maximum Usefulness

Jiang found that the differential privacy framework is actually too powerful in many respects, particularly in medical studies for rare diseases where the number of participants may be fewer than 1,000. In these cases, it is paramount to protect subjects' identities while still drawing scientifically valid conclusions. "The noise added to the data set [for differential privacy] is usually huge," she said. "This means that you may not get the same inference results by using the synthetic data set."

Instead, her team chose a model that preserves the statistical conclusions between the raw data and the synthetic data set [2]. Their method is based on the multiple imputation (MI) framework that statistician Donald Rubin originally proposed in the 1980s to draw inferences from data sets with missing entries [3]. Jiang and others turned this framework inside out; they began with existing complete data, built a model for it, then utilized MI to generate synthetic data (see Figure 2, on page 1). "MI is a missing data framework, but we don't have missing data in the sense that the data provider actually has access to the private data set," Jiang explained. "Our framework takes this fact into account, and then we generate additional data based on private data. Because the model is always correctly specified, we can maintain information in our synthetic data. This is the novelty of our approach."

The synthetic data set has the exact same statistical characteristics as the original data because of its construction process. But Jiang warned that this method is most valuable for small studies; more data begets a greater possibility for users to identify correlations between variables, even when the data themselves are artificial. The next phase of research involves circumventing this limitation.

### Testing the Defenses

Governments around the world (including in Canada, where both Jiang and Gambs are based) increasingly require their funded research projects to publish all data. This demand is laudable in general terms, as openness aids replicability and trust in science. Yet in addition to the general need to refrain from exposing private information like names and addresses, researchers also do not want to inadvertently hurt participants in other ways. In particular, multiple forms of discrimination are legal — some U.S. states allow companies to fire LGBTQ+ employees, for example. Many scientific studies need to account for drug usage, sex work, and other widespread activities, and subjects rightfully hesitate to participate if they know that they are at risk of exposure. Employers too often skirt the law even when discrimination is illegal, such as for disability, pregnancy, or chronic illness. These actions make participant protection even more crucial during data publication.

With these concerns in mind, Jiang, Gambs, and their collaborators are investigating ways to prevent both inadvertent exposure and malicious privacy attacks. Regardless of which method is best for the research at hand, the goal remains the same: do not hurt your subjects to obtain a scientific conclusion.

[2] In a previous *SIAM News* article, Matthew Francis wrote about differential privacy and the U.S. census: https://sinews.siam.org/Details-Page/using-differential-privacy-to-protect-the-united-states-census.

### References

[1] Gambs, S., Ladouceur, F., Laurent, A., & Roy-Gaumond, A. (2021). Growing synthetic data through differentially-private vine copulas. *Proc. Priv. Enh. Technol.*, *2021*(3), 122-141.

[2] Jiang, B., Raftery, A.E., Steele, R.J., & Wang, N. (2022). Balancing inferential integrity and disclosure risk via model targeted masking and multiple imputation. *J. Am. Stat. Assoc.*, *117*(537), 52-66.

[3] Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys*. Hoboken, NJ: Wiley & Sons.

[4] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, *8*, 229-231.

*Matthew R. Francis is a physicist, science writer, public speaker, educator, and frequent wearer of jaunty hats. His website is BowlerHatScience.org.*

## Mathematical Paradoxes

*Continued from page 1*

As we gain a deeper appreciation of AI's limitations, we can better understand its foundations and acquire a stronger sense of direction for exciting new AI techniques. This is precisely the type of growth that happened with the work of Gödel and Turing, which respectively lead to many modern foundations and modern computer science.

By expanding the methodologies of Gödel and Turing, we initiate a foundations program about the boundaries of AI and demonstrate limitations on the existence of randomized algorithms for NN training [4]. Despite many results that establish the existence of NNs with excellent approximation properties, algorithms that can compute these NNs only exist in specific cases.

### Desirable NNs May Exist

Classical approximation theorems show that NNs can approximate a continuous function arbitrarily well [5]. In response, we might initially expect few restrictions on the scientific problems that NNs can handle. For example, consider the least absolute shrinkage and selection operator (LASSO) problem

$$\Xi(y) = \operatorname{argmin}_{x\in\mathbb{C}^N} \lambda\|x\|_{l^1} + \|Ax-y\|_{l^2}^2, \quad \lambda > 0 \qquad (1)$$

for a fixed $A\in\mathbb{R}^{m\times N}$ with variable $y\in\mathbb{R}^m$. Can we train a NN to solve this problem? Let us consider a simple scenario wherein we have a collection $\mathcal{S}=\{y_k\}_{k=1}^R$ and want to compute a NN $\varphi:\mathbb{R}^m\to\mathbb{R}^N$, such that $\operatorname{dist}(\varphi(y_k),\Xi(y_k))\le\epsilon$ for some accuracy parameter $\epsilon>0$ and any $y_k\in\mathcal{S}$.

Here, $\operatorname{dist}(x,\Xi(y))$ denotes the $l^2$-distance of $x\in\mathbb{R}^m$ to the solution set $\Xi(y)$. We take the word "compute" literally, meaning that a computer can never exactly give $A$ and the $y_k$s; for example, an entry of $A$ could be an irrational number. Even if $A$ and the $y_k$s are all rational, the overwhelming majority of software runs floating-point arithmetic in base-2. The training data that is available to an algorithm is thus the collection of all $\mathcal{T}=(\{A_n\}_{n\in\mathbb{N}}, \{y_{k,n}\}_{k\le R,n\in\mathbb{N}})$, such that $\|A-A_n\|\le 2^{-n}$ and $\|y_k-y_{k,n}\|\le 2^{-n}$, i.e., an arbitrary precision approximation of the dataset. By denoting the suitable collection of NNs with $\mathcal{NN}$, it follows easily from classical approximation theory that a mapping $\Phi$ exists with

$$\Phi(\mathcal{T})=\varphi_{\mathcal{T}}\in\mathcal{NN}, \quad \text{where} \quad \varphi_{\mathcal{T}}(y)\in\Xi(y) \quad \forall y\in\mathcal{S}. \qquad (2)$$

This formula raises the following question: *If we can prove the existence of a NN with great approximation qualities, can we find the NN with a training algorithm?*

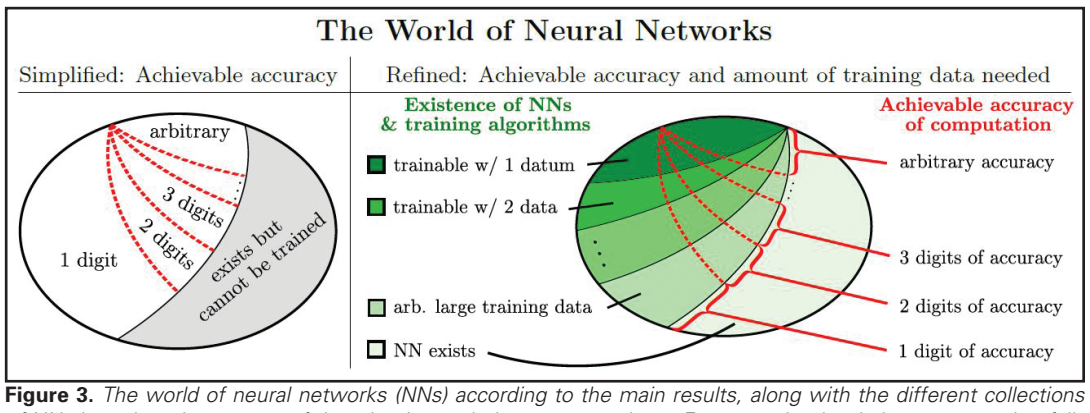



**Figure 3.** *The world of neural networks (NNs) according to the main results, along with the different collections of NNs based on the amount of data that is needed to compute them. For example, the dark green area that falls above the top dashed red line represents the family of NNs that training algorithms can compute to arbitrary levels of accuracy with only one data point. Figure courtesy of the authors.*

### But They May Not Be Trainable

The answer to the aforementioned question is "no," but for quite subtle reasons. Consider the earlier LASSO problem (1). While a NN for this problem provably *exists*—as in (2)—it generally *cannot be trained* by an algorithm [4]. Pick any positive integers $K\ge 3$ and $L$. Well-conditioned classes of datasets, such that (2) is true, do then exist. Yet regardless of computing power and the data's precision levels, we have the following:

**(i) Not trainable:** No algorithm, not even one that is randomized, can produce a NN with $K$ digits of accuracy for any member of the dataset with a probability greater than $1/2$.

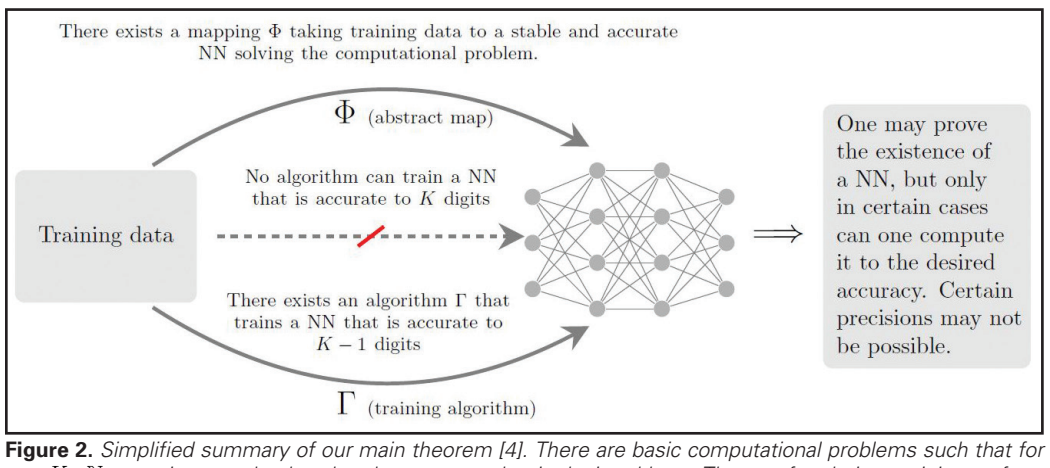



**Figure 2.** *Simplified summary of our main theorem [4]. There are basic computational problems such that for any $K\in\mathbb{N}$, certain cases lead to the phenomenon that is depicted here. The proof technique originates from the mathematics behind the Solvability Complexity Index hierarchy, which generalizes mathematical paradoxes that date back to David Hilbert, Kurt Gödel, and Alan Turing [1, 2, 4, 7,12]. Figure courtesy of the authors.*

**(ii) Not practical:** $K-1$ digits of accuracy is possible over the whole dataset, but any algorithm that trains such a NN requires arbitrarily large training data.

**(iii) Trainable and practical:** $K-2$ digits of accuracy is possible over the whole dataset via an algorithm that only uses $L$ training data, regardless of the parameters.

Figure 3 depicts a Venn diagram of the intricate world of NNs that is based on the above results. We try to compute the existing accurate NN in Figure 4, even though we know that doing so is impossible.

### The SCI Hierarchy

The techniques that prove our results stem from the seminal work of Gödel and Turing, with generalizations and extensions from the Solvability Complexity Index (SCI) hierarchy [2]. The SCI hierarchy and its accompanying tools allow users to obtain sharp boundaries of algorithms' abilities. We expand upon and refine some of the tools that are associated with this hierarchy, as well as the mathematics behind Smale's extended 9th problem about linear programs [1, 11]. To prove our results, we also introduce and develop the theory of *sequential general algorithms*. General algorithms are a key tool within the SCI hierarchy, and sequential general algorithms broaden this concept and capture the notion of adaptive and/or probabilistic choices of training data.

### The Boundaries of AI Through Numerical Analysis

Any theory seeking to understand the foundations of AI must be aware of methodological limitations. This realization is increasingly apparent. "2021 was the year in which the wonders of artificial intelligence stopped being a story," Eliza Strickland wrote in *IEEE Spectrum*.[3] "Many of this year's top articles grappled with the limits of deep learning (today's dominant strand of AI) and spotlighted researchers seeking new paths."

Given the rich history of establishing methodological boundaries via condition numbers, backward errors, precision analysis, and so forth, it is natural to turn to numerical analysis for a solution. We must design a program about the boundaries of AI through numerical analysis to determine the areas wherein modern AI can be made robust, secure, accurate, and ultimately trustworthy. Due to methodological boundaries, such a program cannot include all areas. The formidable question is thus: *When can modern AI techniques provide adequate robustness and trustworthiness?* The answer to this query will shape political and legal decision making and significantly impact the market for AI technologies.

Moreover, we cannot determine this theory solely with the extensive collection of non-constructive existence theorems for NNs, as evidenced by the previous impossibility result. The big challenge is identifying the NNs that are not only stable and accurate, but can also be computed by algorithms. This collection is a small subset of the collection of NNs that are proven to exist.

### References

[1] Bastounis, A., Hansen, A.C., & Vlačić, V. (2021). The extended Smale's 9th problem – On computational barriers and paradoxes in estimation, regularisation, computer-assisted proofs and learning. Preprint, *arXiv:2110.15734*.

[2] Ben-Artzi, J., Colbrook, M.J., Hansen, A.C., Nevanlinna, O., & Seidel, M. (2020). Computing spectra – On the solvability complexity index hierarchy and towers of algorithms. Preprint, *arXiv:1508.03280*.

[3] Choi, C.Q. (2021, September 21). 7 revealing ways AIs fail. *IEEE Spectrum*. Retrieved from https://spectrum.ieee.org/ai-failures.

[4] Colbrook, M.J., Antun, V., & Hansen, A.C. (2022). The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem. *Proc. Nat. Acad. Sci.*, *119*(12), e2107151119.

[5] DeVore, R., Hanin, B., & Petrova, G. (2021). Neural network approximation. *Acta Numer.*, *30*, 327-444.

[6] Finlayson, S.G., Bowers, J.D., Ito, J., Zittrain, J.L., Beam, A.L., & Kohane, I.S. (2019). Adversarial attacks on medical machine learning. *Science*, *363*(6433), 1287-1289.

[7] Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte Math. Phys*, *38*(1), 173-198.

[8] Heaven, D. (2019). Why deep-learning AIs are so easy to fool. *Nature*, *574*(7777), 163-166.

[9] Higham, N.J. (2002). *Accuracy and stability of numerical algorithms*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

[10] Muckley, M.J., Riemenschneider, B., Radmanesh, A., Kim, S., Jeong, G., Ko, J., … Knoll, F. (2021). Results of the 2020 fastMRI challenge for machine learning MR image reconstruction. *IEEE Trans. Med. Imaging*, *40*(9), 2306-2317.

[11] Smale, S. (1998). Mathematical problems for the next century. *Math. Intell.*, *20*, 7-15.

[12] Turing, A.M. (1937). On computable numbers, with an application to the Entscheidungsproblem. *Proc. London Math. Soc.*, *s2-42*(1), 230-265.

*Vegard Antun is a postdoctoral fellow in applied mathematics at the University of Oslo. His research is centered on deep learning-based techniques for scientific computing, with a particular focus on inverse problems and imaging. Matthew J. Colbrook is an FSMP Fellow at École Normale Supérieure in Paris and a Junior Research Fellow at Trinity College in Cambridge. He received the 2022 SIAM Richard C. DiPrima Prize for his Ph.D. thesis and currently focuses on artificial intelligence (AI), spectral computations, and partial differential equations. Anders C. Hansen is a Royal Society University Research Fellow and professor of mathematics at the University of Cambridge, where he leads the Applied Functional and Harmonic Analysis Group. He also focuses on foundations of computations with applications to AI.*

[3] https://spectrum.ieee.org/artificial-intelligence-2021

| $\min_k \operatorname{dist}(\Psi_n(y_k),\Xi(A,y_k))$ | $\min_k \operatorname{dist}(\Phi_n(y_k),\Xi(A,y_k))$ | prec. of training data | $10^{-K}$ |
|---|---|---|---|
| 0.2999690 | 0.2597827 | $n=10$ | $10^{-1}$ |
| 0.3000000 | 0.2598050 | $n=20$ | $10^{-1}$ |
| 0.3000000 | 0.2598052 | $n=30$ | $10^{-1}$ |
| 0.0030000 | 0.0025980 | $n=10$ | $10^{-3}$ |
| 0.0030000 | 0.0025980 | $n=20$ | $10^{-3}$ |
| 0.0030000 | 0.0025980 | $n=30$ | $10^{-3}$ |
| 0.0000030 | 0.0000015 | $n=10$ | $10^{-6}$ |
| 0.0000030 | 0.0000015 | $n=20$ | $10^{-6}$ |
| 0.0000030 | 0.0000015 | $n=30$ | $10^{-6}$ |

**Figure 4.** *Impossibility of computing approximations of the neural network (NN) to arbitrary accuracy. We demonstrate the impossibility statement on fast iterative restarted networks $\Phi_n$ and learned iterative shrinkage thresholding algorithm networks $\Psi_n$ [4]. The table reveals the shortest $l^2$-distance between the networks' output and the problem's true solution for different values of $n$ (precision of training data is $2^{-n}$) and $K$ (integer from the theorem). Neither of the trained NNs can compute the existing correct NN to $10^{-K}$ digits of accuracy, but both compute approximations that are accurate to $10^{-K+1}$ digits. Figure courtesy of [4].*

## Take Advantage of SIAM's Visiting Lecturer Program

Hearing directly from working professionals about research, career opportunities, and general professional development can help students better understand the workforce. SIAM facilitates such interactions through its Visiting Lecturer Program (VLP), which is sponsored by the SIAM Education Committee and offers a roster of experienced applied mathematicians and computational scientists in industry, government, and academia. Mathematical sciences students and faculty—including SIAM student chapters—can invite VLP speakers to present about topics that are of interest to developing professional mathematicians. Why not host a SIAM visiting lecturer for a virtual talk?

Points to consider when deciding to host a visiting lecturer include the choice of dates; speakers; topics; and any additional or related activities, such as follow-up discussions. Organizers can reach out directly to speakers and must address these points when communicating with them. Read more and view the current list of speakers online.[1]

[1] https://www.siam.org/students-education/programs-initiatives/siam-visiting-lecturer-program

## Biological Network

*Continued from page 2*

of evolution in biology. Here we provide an example of each of these aspects.

**The organizational principles of the networks themselves can be different.** It is well known that social networks obey the principle of "triadic closure" so that triangles are likely; i.e., a friend of many of my friends is also likely to be my friend [9]. When we perform link prediction on a social network, some of the highest confidence predicted links tend to occur between nodes with a lot of common neighbors. However, many protein-protein interactions are of a "lock and key" type. Specifically, if a set of keys exists for which a large overlapping subset opens two locks, we should predict that the rest of the keys that open lock 1 will also open lock 2, not that the two locks directly interact. This concept forms the basis of the recently proposed "length three" measure [8]; incorporating this insight into a more complicated embedding measure improves PPI link prediction methods [7].

**The nature and availability of ground-truth data to test methods may differ.** As remarked previously, the disease-gene module detection problem for biological networks is directly analogous to the so-called unsupervised community detection problem for social networks. However, method performance for community detection in social networks is usually measured in one of two settings:

(i) A setting in which ground-truth communities are given (either from real data sets or via synthetic data wherein a clique or dense subgraph is planted [2]), and we measure how well the method recovers these communities.

(ii) A setting in which we assume that no ground-truth communities are known and instead measure performance based on the clusters' mathematical coherence, according to standards such as modularity or conductance.

Neither setting perfectly captures the typical situation in computational biology domains, where ground truth availability is rarely all or nothing. Rather, we are most often in an in-between state with partial ground truth—not uniformly sampled—from which we can nonetheless still derive principled performance tests based on how well our method aligns with the available data. For example, the designers of the 2016 DREAM Challenge conceived a clever way to test the predicted communities' alignment with known biological disease genes. They used curated sets of genome-wide association study sequencing data as an independent empirical source to produce sets of genes that were associated with human disease, then awarded points for each community in which more genes for a certain disease occurred together than otherwise expected [4].

**We can leverage the power of evolution.** Many experiments that attempt to understand human disease are performed on model organisms like yeast, fruit flies, or mice. Because some of this information is encoded in the form of a PPI network, a powerful idea involves the use of co-embedding to simultaneously embed both networks in a space where local neighborhoods indicate functional similarity, even across species. On the surface, this looks a lot like network alignment [11]: another well-studied problem in the context of social networks. But if the two species are as evolutionary distant as humans and mice, we cannot expect a one-to-one alignment. While we can match some human genes to unique mouse genes that evolved from a common ancestor, the mapping for other genes may be many to one or many to none if the gene in question was duplicated or lost in one or both of the species. We therefore examine methods that perform a bijective alignment of a subset of genes with unique matches, then utilize distance within each network to complete the embedding. We unlock improvements in function prediction for human genes as well as for less-studied species of interest with minimal experimental data [1].

### Outlook

The rich and important set of problems within the space of biological networks yields a fruitful path to important collaborations between experts in computational biology and network science. As we customize and develop a toolbox for the mathematical analysis of biological networks, critical biological and biomedical applications will continue to emerge.

*This article is based on Lenore Cowen's invited presentation at the 2021 SIAM Conference on Applied and Computational Discrete Algorithms,[2] which took place virtually last year in conjunction with the 2021 SIAM Annual Meeting.[3]*

### References


[1] Arsenescu, V., Devkota, K., Erden, M., Shpilker, P., Werenski, M., & Cowen, L.J. (2022). MUNDO: Protein function prediction embedded in a multispecies world. *Bioinform. Adv.*, *2*(1), vbab025.

[2] Bickel, P.J., & Chen, A. (2009). A nonparametric view of network models and Newman-Girvan and other modularities. *Proc. Nat. Acad. Sci.*, *106*(50), 21068-21073.

[3] Cao, M., Zhang, H., Park, J., Daniels, N.M., Crovella, M.E., Cowen, L.J., & Hescott, B. (2013). Going the distance for protein function prediction: A new distance metric for protein interaction networks. *PloS One*, *8*(10), e76339.

[4] Choobdar, S., Ahsen, M.E., Crawford, J., Tomasoni, M., Fang, T., Lamparter, D., … Marbach, D. (2019). Assessment of network module identification across complex diseases. *Nat. Methods*, *16*(9), 843-852.

[5] Cowen, L., Devkota, K., Hu, X., Murphy, J.M., & Wu, K. (2021). Diffusion state distances: Multitemporal analysis, fast algorithms, and applications to biological networks. *SIAM J. Math. Data Sci.*, *3*(1), 142-170.

[6] Cowen, L., Ideker, T., Raphael, B.J., & Sharan, R. (2017). Network propagation: A universal amplifier of genetic associations. *Nat. Rev. Genet.*, *18*(9), 551-562.

[7] Devkota, K., Murphy, J.M., & Cowen, L.J. (2020). GLIDE: Combining local methods and diffusion state embeddings to predict missing interactions in biological networks. *Bioinform.*, *36*(Supplement 1), i464-i473.

[8] Kovács, I.A., Luck, K., Spirohn, K., Wang, Y., Pollis, C., Schlabach, S., … Barabási, A.-L. (2019). Network-based prediction of protein interactions. *Nat. Comm.*, *10*(1), 1240.

[9] Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks. *J. Assoc. Inform. Sci. Tech.*, *58*(7), 1019-1031.

[10] McPherson, M., Smith-Lovin, L., & Cook, J.M. (2001). Birds of a feather: Homophily in social networks. *Ann. Rev. Sociol.*, *27*(1), 415-444.

[11] Singh, R., Xu, J., & Berger, B. (2008). Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proc. Nat. Acad. Sci.*, *105*(35), 12763-12768.

[12] Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., ... von Mering, C. (2021). The STRING database in 2021: Customizable protein-protein networks and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.*, *49*(D1), D605-D612.


*Lenore J. Cowen is a professor of computer science at Tufts University, with joint appointments in the Department of Mathematics and the Genetics Program. She currently serves as principal investigator and director of the Tufts T-TRIPODS Institute[4] for foundational data science.*

[2] https://www.siam.org/conferences/cm/conference/acda21

[3] https://www.siam.org/conferences/cm/conference/an21

[4] https://tripods.tufts.edu

# John von Neumann: The Man From the Future

**The Man from the Future: The Visionary Life of John von Neumann.** *By Ananyo Bhattacharya. W.W. Norton & Company, New York, NY, February 2022. 368 pages, $30.00.*

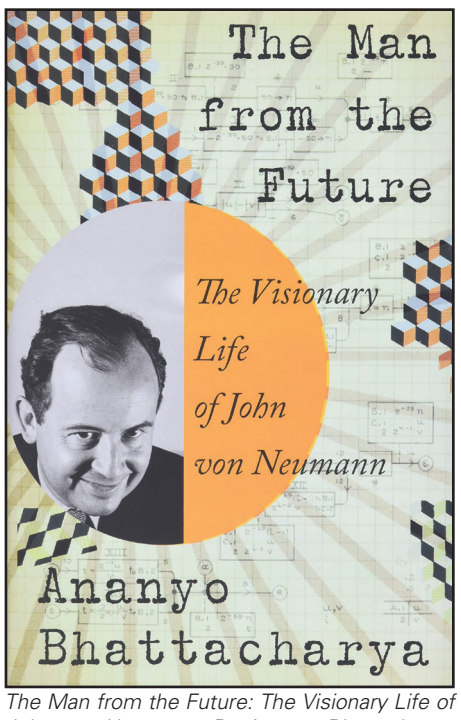
BOOK REVIEW
*By Paul Davis*

Members of the SIAM community may come to Ananyo Bhattacharya's fresh and lively biography of John von Neumann—*The Man from the Future*—with a predictably positive bias. SIAM's highest professional honor is the John von Neumann Prize,[1] which is awarded for "distinguished contributions to the field of applied mathematical sciences and for the effective communication of these ideas to the community." It was established to honor the Hungarian-American mathematician, physicist, and computer scientist whose seminal work helped found the field of modern computing. Winners are recognized at the SIAM Annual Meeting, where they deliver an associated flagship lecture; Leah Edelstein-Keshet of the University of British Columbia is the 2022 prize recipient.[2]

For many of us, the foundations of our own research have at least one brick that has been shaped by von Neumann's ideas. Bhattacharya's account of von Neumann's life, research, and varied collaborations will refine readers' understanding of his impact on the larger scientific and social milieu in which he worked, as well as the continuing influence of his legacy long after his death.

Bhattacharya naturally begins with von Neumann's birth in 1903 and his childhood among the elite in "sparkling Belle Epoch Budapest." He honed a formidable intelligence among schoolmates like future economist William Fellner and physicist Eugene Wigner, in addition to older mentors such as Theodore von Kármán, a pioneer of fluid dynamics.

While preparing to leave high school (more precisely, "real school" — a variant of the European *gymnasium*) at the age of 17, von Neumann began his first major mathematical work. In a bold, early draft of what eventually became his doctoral thesis, he sought "to make Cantor's ordinal numbers unambiguous and concrete." The necessity of this task arose from the logical turmoil of Bertrand Russell's famous set-theoretic paradox: Can the set of all sets that are not members of themselves be a member of itself? Russell's paradox had materialized among the fallout from David Hilbert's challenge to place all mathematics on an unshakable axiomatic foundation. The concepts of ordinality and cardinality that von Neumann developed in response to this challenge are still in use today.



*The Man from the Future: The Visionary Life of John von Neumann. By Ananyo Bhattacharya. Courtesy of W.W. Norton & Company.*

Five years later and with a doctorate in hand (as well as a spare chemical engineering degree to relieve his father's concerns about future employment), von Neumann followed a grant from the Rockefeller Foundation to Hilbert's department in Göttingen — the center of the mathematical universe at the time. He stepped directly into another intellectual maelstrom: Werner Heisenberg's matrix formulation of quantum mechanics was butting against Erwin Schrödinger's wave theory amid no end of physically implausible behavior. Bhattacharya provides a vivid and accessible play-by-play account of the intellectual turmoil that von Neumann eased by deploying Hilbert's spectral theory to reconcile the two apparently different mathematical formulations and begin untangling the physical conundrums.

The author's rich account of quantum disputes among physics luminaries continues as he follows von Neumann to the University of Berlin, then to the University of Hamburg. In Berlin, the 23-year-old *Privatdocent*—the youngest ever appointed at the time—reveled in the city's decadent post-war atmosphere. His childhood friend Wigner happened to be in Berlin at the same time, living as a scholarly hermit who only emerged to attend physics colloquia. He reported that von Neumann "was sort of a bon vivant, and went to cabarets and all that."

In 1929, von Neumann moved briefly to Hamburg before receiving an invitation for a highly paid lectureship at Princeton University. He engineered a parallel invitation for Wigner, and the two arrived in the U.S. nearly simultaneously. Von Neumann was accompanied by both his new bride, Mariette Kövesi, and a prodigious reputation based on his foundational contributions to axiomatic set theory and quantum mechanics.

Wigner recounted that at their first meeting in America, he and von Neumann quickly "agreed that we should try to become somewhat American." According to Wigner, the Johnny von Neumann who emerged "was a cheerful man, an optimist who loved money and believed firmly in human progress."

With von Neumann in the U.S., Bhattacharya's illuminating sketches of the personalities and events of early and mid-century European mathematics and physics expand to include America and the emergence of applied and computational mathematics during and after World War II. At this point, readers might feel as if they have stumbled into the story of a fictional time traveler from the *Outlander* series. Von Neumann somehow managed to be on location for every major mathematical happening; he was in Budapest long enough to start repairing an essential part of the foundations of mathematics, then moved to Göttingen just as Schrödinger's cat began to prowl.

Once von Neumann crossed the Atlantic and plunged into America's war efforts, his pace quickened. The mathematical traveler was in Los Alamos to compute shock waves within bomb detonators and dispassionately select targets for atomic bombs; in Aberdeen to automate the calculation of ballistics tables; in Princeton, Santa Monica, and elsewhere to develop game theory; and almost everywhere—Philadelphia, Poughkeepsie, and maybe even London for a surreptitious collaboration (who knows?) with Alan Turing—to crystallize the operational and design principles of stored-program computers.

Perhaps Von Neumann's most audacious stop was nominally in Pasadena, where he delivered the first major public presentation of his ideas about self-replicating automata. His abstract model was a universal Turing machine, and his thinking lay well ahead of what biologists knew about DNA and the mechanics of coding, copying, and replicating life forms. The entire conception seemed to draw science perilously close to science fiction — and to the future.

Throughout his book, the author's descriptions of the surrounding scientific landscape are generally so deft and effective that readers may feel that they are learning more about the setting than the protagonist. On the other hand, Bhattacharya frequently blends the reach of the scientific challenges that intrigued von Neumann with accounts of his achievements in ways that ultimately provide shadowy silhouettes of both his character and his scientific powers. For readers of a certain age, this chronicle offers some of the same pleasures of reminiscing with an old family album and connecting legendary names, places, and events from previous generations to our own. It also reminds us of Hungary's disproportionate gifts to 20th-century mathematics.

*See* **John von Neumann** *on page 8*

[1] https://www.siam.org/prizes-recognition/major-prizes-lectures/detail/john-von-neumann-prize

[2] https://sinews.siam.org/Details-Page/leah-edelstein-keshet-is-the-2022-siam-john-von-neumann-prize-lecturer

## Novel Ph.D. Program

*Continued from page 3*

sense of learning how to set up and build models in the first place.

There is also one component that is fairly unique, especially within mathematics programs: we require all students to complete an interdisciplinary internship. We keep things pretty flexible; our minimum requirement is a paid, full-time position that is outside of the university and at least a term long — either a full semester or eight-10 weeks during the summer. We've had students work in pharmaceutical companies, research wings of hospitals, and consulting companies. Some have done great work in national labs or government organizations, including Los Alamos National Laboratory[3] (LANL), Oak Ridge National Laboratory,[4] and the Air Force Research Laboratory.[5] We even allow students to conduct research in the labs of potential experimental collaborators at other universities. It's been a great experience for them, and many have received offers for further internships.

**SN: What sort of qualifications do you look for in prospective students?**

**NC:** Applicants don't have to be applied math majors to find success in our program. Instead, our admissions committee looks for a specific interest in mathematical modeling and strong evidence of success in undergraduate coursework that is related to linear algebra, probability, differential equations, and the like. We also see whether prospective students have demonstrated that they're not afraid to get their hands dirty and do some coding. A lot of our applicants have strong research experience, either with faculty members or through the National Science Foundation's Research Experiences for Undergraduates program,[6] and that's great. However, we recognize that some students don't have the opportunity to spend a summer doing research.

As we grow, we hope to attract undergraduates who realize that their passion is not just *doing* mathematics, but specifically finding ways to apply mathematics to complicated, real-world problems.

**SN: What types of careers do students envision for themselves after graduation?**

**NC:** A number of students have found real interest in the government sector, and some have gone to the national labs. Two students recently received fellowships through the Oak Ridge Institute for Science and Education.[7] Both are working with the U.S. Food and Drug Administration to validate and analyze COVID-19 models and data.

A lot of students are interested in biomedical applications and could work for companies that are devoted to health care and/or pharmaceuticals. And a number are trying to discern whether they want to follow an academic path after realizing that many research opportunities do exist in academia.

**SN: Can you share a few examples of student success stories?**

**NC:** Nicole Rosato, our first graduate, currently works at Rochester Gas and Electric. The title of her job is "data engineer," but she does a great deal of data science. She develops models to analyze and predict outage events — causes of outages, when and where they're going to occur, potential impacts, and possible fixes to prevent them from happening in the first place.

I'd also like to highlight Jenna McDanold, who came to our program as more of a nontraditional student. She is an artist who specializes in woodburning, and she wanted to learn how to apply mathematical modeling to wildfire spread. During her second year, Jenna reached out to a group at LANL under Rodman Linn,[8] one of the world's leading wildfire experts. She completed an internship there and they liked her so much that they asked her to stay. Now she's a research assistant at LANL, which is supporting her as she finishes her Ph.D.

**SN: What do you envision for the program's future?**

**NC:** RIT faculty members Kara Maki and David Ross recently founded the Industrial Math Modeling Center[9] (IMMC) within the College of Science at RIT, which started with connections that they had with local companies. Kara and David envision the IMMC as a place where organizations can connect with mathematical modeling students as prospective interns or employees. We hope that the IMMC will be integrally associated with our Ph.D. program; as more companies feed us industrial problems in the next five to 10 years, we will connect students and faculty members to these projects so that they lead to internships and turn into dissertation topics.

**SN: What do you hope that students will take away from this program?**

**NC:** We want students to feel comfortable entering a new environment or application area—a company, government lab, research setting in academia, or even a hospital system—where they don't necessarily have much background knowledge. If they can talk with experts in that area, understand the relevant problems, and use mathematical modeling to propose techniques to solve these problems, then we'll have done our job. We're trying to help students build these foundational skills because new and complicated issues always arise in the world. 10 years from now, there will be problems that we can't even envision — problems that people with strong analytical skills and mathematical backgrounds will be able to tackle. If we produce graduates who are capable of entering new settings and communicating with experts in those fields, then our program will have been wildly successful.

*Lina Sorg is the managing editor of SIAM News.*

[3] https://www.lanl.gov

[4] https://www.ornl.gov

[5] https://www.afrl.af.mil

[6] https://www.nsf.gov/crssprgm/reu

[7] https://orise.orau.gov

[8] https://www.lanl.gov/search-capabilities/profiles/rod-linn.shtml

[9] https://www.rit.edu/science/industrial-math-modeling-center

# SIAM and John von Neumann

*By Paul Davis*

SIAM's John von Neumann Prize[1] is naturally the best-known connection between SIAM and the esteemed mathematician John von Neumann. It was established in 1959 as the "von Neumann Lecture," with initial monetary contributions from IBM and other organizations. Once those assets were exhausted, SIAM funded all of the award's expenses and continues to do so.

"Initially, the lecture alternated between someone inside the applied math community and someone from a more applied discipline that used a lot of applied math," James Crowley, the former executive director of SIAM, explained. "You hence see, for example, some notable economists in the list of former lecturers. Over time, the leadership began to regard this lecture as a flagship prize of SIAM, and the focus changed from alternating between applied mathematicians and users of applied math to recognizing the best in our field." A few years ago, the Board of Trustees formally recognized this evolution by changing the designation from *Lecture* to *Prize* and noting that it is the highest honor that SIAM bestows. The full list of past recipients is available on the prize webpage.[2]

The prize signifies just one of many connections between von Neumann and the SIAM membership. At the 2003 SIAM Conference on Computational Science and Engineering,[3] Gene Golub and Joseph Grcar organized "John von Neumann's 100th Birthday Celebration Symposium." Philip J. Davis provided a personal overview of the celebration in a *SIAM News* article[4] [2] and recounted the remarks of the four speakers: William Aspray[5] [1], Marina von Neumann Whitman[6] [6], Peter Lax[7] [4], and Pete Stewart (the first three talks subsequently appeared in *SIAM News* in 2005). The presentations collectively served as a rich mix of accounts and anecdotes that captured von Neumann's personal and mathematical influences throughout the history of computing.

*John von Neumann, 1903-1957. Photo courtesy of the U.S. Department of Energy.*

In addition, Grcar has analyzed in illuminating detail [3] what he calls "the first modern paper in numerical analysis" — von Neumann and Herman Goldstine's study of Gaussian elimination with ideas that are now well known under their modern names, such as condition number and Courant-Friedrichs-Lewy stability criteria [5]. Of course, this research was the initial gust in a mighty storm of rigorous work in numerical computing, much of which was led by SIAM members and published in SIAM journals — notably the *SIAM Journal on Numerical Analysis*.[8]

In his paper, Grcar also recounts many specifics about the intellectual property dispute that arose in part from Goldstine's circulation of von Neumann's "First Draft of a Report on the EDVAC."[9] This obviously incomplete text was the first description of an architecture for a stored-program computer, with only one reference and von Neumann as its single author. It was written during a time when von Neumann would have been fully cognizant of J. Presper Eckert and John Mauchly's work with ENIAC — an automatic calculator and predecessor to the programmable EDVAC.

Ananyo Bhattacharya's new von Neumann biography, *The Man from the Future*,[10] offers further details and a slightly different perspective. However, the key takeaway is that von Neumann's report became the legal cannon that crippled Eckert and Mauchly's plans to commercialize their work. The duo's firm did ultimately produce the UNIVAC I (UNIVersal Automatic Computer I), a machine that was sold at a loss to the U.S. Department of Commerce to automate the analysis of the 1950 census. In the meantime, von Neumann enjoyed a well-paid consulting assignment with IBM while the patent disputes, the historic question of "Who invented the computer?," and the dilemmas of public and private rights simmered on for years.

Mauchly went on to serve as the fourth president of SIAM — a critical contribution in SIAM's embryonic, volunteer-driven early years. He is now largely overlooked in the shadows of von Neumann's larger and well-deserved reputation.

A different and broader cache of stories from the youthful years of scientific computing is available in SIAM's "History of Numerical Analysis and Scientific Computing,"[11] which contains fascinating oral histories, articles, and other resources about key players and developments. While many of the modern chapters began with von Neumann and Goldstine's famous Gaussian elimination paper [5], the story grows richer and more satisfying every day. And SIAM remains the place to follow it.

## References


[1] Aspray, W. (2005, March 1). On John von Neumann and his role in the development of the computer. *SIAM News*, *38*(2), p. 7.

[2] Davis, P.J. (2003, May 30). SIAM conference on CS&E: John von Neumann at 100: SIAM celebrates a rich legacy. *SIAM News*, *36*(4), p. 1.

[3] Grcar, J.F. (2011). John von Neumann's analysis of Gaussian elimination and the origins of modern numerical analysis. *SIAM Rev.*, *53*(4), 607-682.

[4] Lax, P. (2005, March 1). John von Neumann: The early years, the years at Los Alamos, and the road to computing. *SIAM News*, *38*(2), p. 9.

[5] Von Neumann, J., & Goldstine, H.H. (1947). Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc.*, *53*, 1021-1099.

[6] Von Neumann Whitman, M. (2005, March 1). Life with father. *SIAM News*, *38*(2), p. 8.


*Paul Davis is professor emeritus of mathematical sciences at Worcester Polytechnic Institute.*

[1] https://www.siam.org/prizes-recognition/major-prizes-lectures/detail/full-prize-specifications/john-von-neumann-prize

[2] https://www.siam.org/prizes-recognition/major-prizes-lectures/detail/john-von-neumann-prize

[3] https://archive.siam.org/meetings/cse03/index.htm

[4] https://go.siam.org/ZDgP1d

[5] https://go.siam.org/8mkZ9L

[6] https://go.siam.org/XZ2YSN

[7] https://go.siam.org/htwbWm

[8] https://www.siam.org/publications/journals/siam-journal-on-numerical-analysis-sinum

[9] https://web.mit.edu/STS.035/www/PDFs/edvac.pdf

[10] See page 6 for Paul Davis' review of *The Man from the Future: The Visionary Life of John von Neumann.*

[11] http://history.siam.org

# SIAM Federal Research Priorities Advance as Congress Passes Omnibus Spending Package

*By Andrew Herrin and Miriam Quintal*

After nearly six months of stop-gap funding measures that maintained government spending at fiscal year (FY) 2021 levels, Congress passed a spending package on March 11 for the remainder of FY 2022. This package follows President Biden's first year in office, during which the administration set ambitious funding goals; pursued massive legislative packages to improve the nation's infrastructure; and addressed public health, climate and clean energy, innovation, education, and racial equity — all while continuing to combat the COVID-19 pandemic. Throughout the first year of the Biden administration, the SIAM Committee on Science Policy[1] has championed increased funding for federal agencies that support applied mathematics and computational science.

The FY 2022 omnibus package includes moderate increases for agencies such as the National Science Foundation (NSF), Department of Energy (DOE), Department of Defense (DOD), and National Institutes of Health (NIH). The increases will allow these agencies to move forward with major initiatives like the new Advanced Research Projects Agency for Health[2] (ARPA-H); a new NSF Directorate for Technology, Innovation, and Partnerships (TIP); and many climate and infrastructure programs.

The omnibus will fund the NSF at $8.838 billion, an increase of $351.24 million (or 4.1 percent) from the FY 2021 enacted level. Though this represents the largest growth in NSF funding in more than a decade, the funding level remains far below the major increases that were proposed in President Biden's budget request[3] and House and Senate draft bills.

The Research and Related Activities (R&RA) account—which includes the Division of Mathematical Sciences and all NSF research directorates—will be funded at $7.16 billion, an increase of $249.63 million (or 3.6 percent) from the FY 2021 level. The omnibus also approves the creation of the new TIP Directorate within R&RA but does not specify a funding level. The explanatory statement that accompanies the omnibus outlines TIP's goals[4] "to advance science and engineering research leading to breakthrough technologies, to find solutions to national and societal challenges, to strengthen U.S. global competitiveness, and to provide training opportunities for the development of a diverse STEM workforce." The Directorate for Education and Human Resources—which funds programs that broaden participation in undergraduate and graduate education, among other activities—will receive $1.01 billion, a 3.9 percent increase over FY 2021 funding levels. Undergraduate education programs are flat funded at FY 2021 levels, but programs and graduate fellowships that broaden participation will each see modest increases of four to seven percent.

The omnibus will provide $44.9 billion for DOE: an increase of $5.2 billion (or 13 percent) above the FY 2021 enacted level. Consistent with the Biden administration's priorities to accelerate the development and deployment of clean energy technologies to meet ambitious net zero carbon goals, the omnibus significantly increases investments in all fundamental and applied energy programs. Mathematical, computational, and computer science research will be funded at $260 million, which represents an increase of $10 million (or four percent) from FY 2021 enacted levels. The DOE's Computational Science Graduate Fellowship[5] is set to receive its first increase in many years; funding for the program will reach at least $15 million — a 50 percent increase over the FY 2021 enacted level. Funding for artificial intelligence (AI) and machine learning (ML) across all Office of Science programs faces a small cut to its $120 million budget.

DOD's science and technology accounts—which include basic research, applied research, and advanced technology development—will be funded at $18.8 billion, a 12 percent increase to the FY 2021 level. Basic research accounts across the services and department wide will experience a five percent increase overall, with the Army securing the largest increase (10.4 percent) and the Air Force collecting the smallest increase (0.8 percent). Basic research at the Defense Advanced Research Projects Agency (DARPA) will collect $446 million — $50 million over the budget request but six percent below the FY 2021 funding level. Congress provides specific funding for AI, ML, and cyber efforts within the Army and DARPA. For the first time, Congress will also bestow $15 million to the new Space Force[6] for fundamental research, which could include novel basic research efforts.

The omnibus will supply a total of $45 billion to NIH in FY 2022 — an increase of $2.03 billion (or 4.7 percent) over the FY 2021 enacted level. This marks the seventh consecutive funding increase for NIH. In addition, Congress will provide $1 billion to the Secretary of Health and Human Services (HHS) to establish ARPA-H. The HHS Secretary can choose to transfer this money and is expected to move the funding to NIH. ARPA-H, which was first articulated in President Biden's budget request, aims to speed up transformational innovation in health research; support high-risk, high-reward research; and accelerate the translation of fundamental biomedical research into clinical applications to provide more treatments and cures for disease. SIAM has been engaging with NIH and the White House Office of Science and Technology Policy[7] to encourage ARPA-H's support of innovations in applied mathematics, computational science, data, AI, and ML.

Now that FY 2022 appropriations are finalized, Congress turns to FY 2023 appropriations and awaits the president's FY 2023 budget request. SIAM will continue to advocate for strong funding for applied mathematics and computational science programs at relevant agencies, and will keep members informed as the process unfolds.

*Andrew Herrin is an associate and Miriam Quintal is Managing Principal at Lewis-Burke Associates LLC, SIAM's governmental relations partner. They are SIAM's liaisons in Washington, D.C.*

[1] https://www.siam.org/about-siam/committees/committee-on-science-policy-csp
[2] https://www.nih.gov/arpa-h
[3] https://sinews.siam.org/Details-Page/siam-advocates-for-research-growth-as-biden-administration-releases-funding-request
[4] https://www.aip.org/fyi/2021/fy22-budget-outlook-national-science-foundation
[5] https://science.osti.gov/ascr/CSGF
[6] https://www.spaceforce.mil
[7] https://www.whitehouse.gov/ostp

## John von Neumann

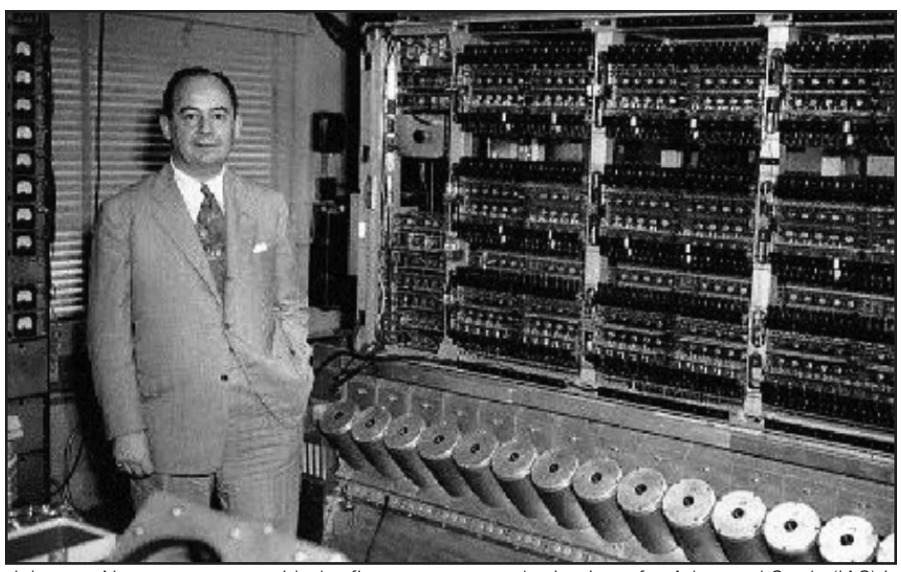
*Continued from page 6*

In addition, these accounts prompt questions—perhaps many are unanswerable—about the man, his life, and his work. Would von Neumann's decisive selection of atomic bomb targets in Japan have been so dispassionate if Germany were still fighting? Would his classic *Theory of Games and Economic Behavior* with economist Oskar Morgenstern have ventured beyond collaborative games if he were not the quintessential centered European? If von Neumann had lived into the 1970s, might he have spoken of the RAND Corporation and its clones in a paraphrase of his famous comment about the relative simplicity of mathematics: "If you think your simple mathematical model explains everything, you don't understand how complicated life is"? What ideas might he pour across a colloquium's video screen today?

Von Neumann may indeed have been *The Man from the Future* who returned to steer his era toward what he had seen. Bhattacharya's sweeping and well-informed biography makes an even better case that von Neumann was a genius of multiple talents who had sensed the future but never seen it fully. His conception of the future was sometimes imperfect, but it was shaped so well that disparate segments of the scientific and technical community could pull it toward realization, shape it to meet new complexities, and ride it to new opportunities. Bhattacharya adroitly describes many of the challenges that occupy applied and computational mathematicians while recounting one of our profession's most important origin stories to a broader audience.[3]

*Paul Davis is professor emeritus of mathematical sciences at Worcester Polytechnic Institute.*

[3] Readers can find more specialized accounts in the accompanying article on page 7, "SIAM and John von Neumann."

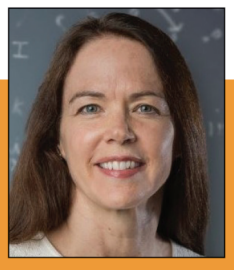


*John von Neumann poses with the first computer at the Institute for Advanced Study (IAS) in Princeton, NJ. Photo courtesy of the Archives of the IAS.*

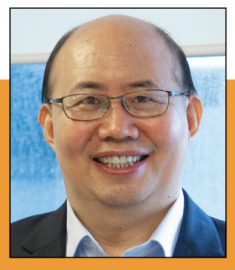
# SIAM Fellows 2022

**SIAM is pleased to announce the newly selected Class of SIAM Fellows—a group of distinguished members of SIAM who were nominated by their peers for exceptional contributions to the fields of applied mathematics and computational science. Please join us in congratulating these 26 members of our community.**

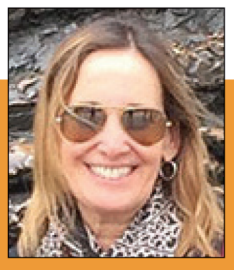
**Remi Abgrall**
Universität Zürich



**Sharon F. Arroyo**
The Boeing Company

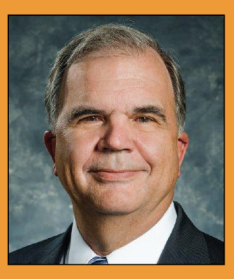


**Weizhu Bao**
National University of Singapore

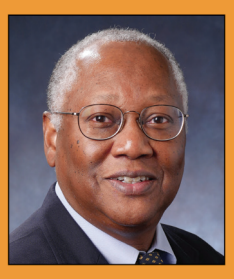


**Bonnie Berger**
Massachusetts Institute of Technology

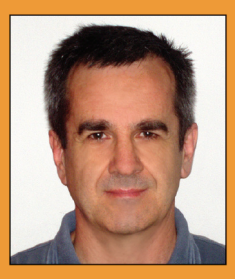
**Zhiming Chen**
Chinese Academy of Sciences



**James Michael Crowley**
SIAM



**James H. Curry**
University of Colorado Boulder



**Zlatko Drmač**
University of Zagreb

**Chen Greif**
The University of British Columbia

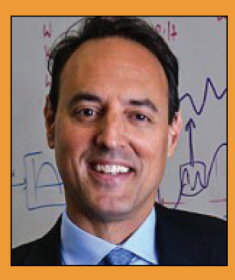
**Abba B. Gumel**
Arizona State University

**Eldad Haber**
The University of British Columbia

**John Robert King**
University of Nottingham

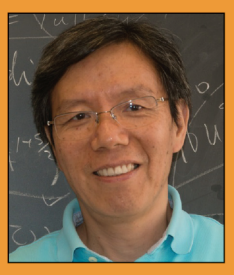
siam Fellow

**Daniel Kressner**
EPFL



**Jose Nathan Kutz**
University of Washington

**Lek-Heng Lim**
University of Chicago



**Fang-Hua Lin**
New York University

**Peter B. Monk**
University of Delaware

**Houman Owhadi**
California Institute of Technology

**Keith Promislow**
Michigan State University

**Rosemary Anne Renaut**
Arizona State University

**Wil Schilders**
Eindhoven University of Technology

**Leonard J. Schulman**
California Institute of Technology

**Amit Singer**
Princeton University

**Gabriele Steidl**
Technische Universität Berlin

**Raymond Tuminaro**
Sandia National Laboratories

**Hongkai Zhao**
Duke University

Learn more about these esteemed individuals and nominate a colleague by October 19, 2022 for the 2023 Class of SIAM Fellows:

**siam.org/Prizes-Recognition/Fellows-Program**

siam | Society for Industrial and Applied Mathematics

# Spinning Tops in Spinning Frames

Spinning tops have existed since ancient times. The popular toy is so old that it is difficult to say something simultaneously (i) new, (ii) correct, and (iii) interesting about it. The Lagrange top is an axisymmetric body that pivots on a needlepoint. When spun quickly enough, it precesses around the vertical axis while also possibly nutating.[1] In most books about classical mechanics [1, 3], the top is treated with somewhat lengthy calculations. Here I describe a point (pun intended) that is lost in standard treatments.

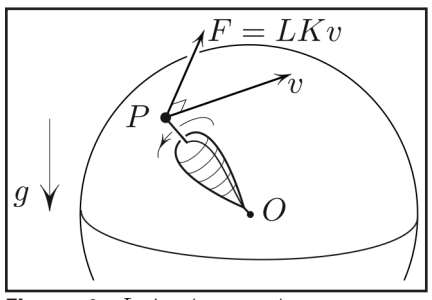



**Figure 1.** *$L$ is the angular momentum around the symmetry axis and $K = R^{-2}$ is the Gaussian curvature of the sphere.*

Let us keep track of the point $P$ where the top's axis punctures a sphere centered at $O$ (see Figure 1). Dynamics of the top admit a remarkably simple description:

*The motion of $P$ is identical to that of a point mass $m$ constrained to the sphere of radius $R$ and subject to two forces: (i) gravity and (ii) the (magnetic-like) force perpendicular to the velocity and of magnitude equal to the product of axial angular momentum $L$, the Gaussian curvature $K = R^{-2}$ of the sphere and the velocity.*

[1] "Nutus" = "nod" in Latin.

Both $R$ and $m$ are specified in terms of the top's mass, moments of inertia, and the center of mass' distance to the pivot; these expressions—as well as the claim's proof—are available in [2]. The Lagrange top is thus equivalent to the charged particle on the sphere in the magnetic field that is perpendicular to the sphere (as if a magnetic monopole was at $O$) and subject to gravity. The strength of the magnetic field is given by the sphere's Gaussian curvature, and the charge is the angular momentum $L$.

## Near-vertical Motions

Figure 2 shows projection $z = (x, y) = x + iy$ of $P$ onto the tangent plane to the north pole. For near-vertical motions, $z$ is a good approximation for $P$ and it satisfies, up to higher order terms in $|z|$,

$$\ddot{z} = ib\dot{z} + az. \qquad (1)$$

Here, $a = g/R$ and $b = LK/m$, although this is not important for the forthcoming point. Figure 3 shows typical trajectories of $z$. Interestingly, when viewed in a rotating frame they are uniform circular motions — provided that we choose the right angular velocity for our frame. Indeed, any solution of (1) is a combination of two circular motions:

$$z = Ae^{i\omega_1 t} + Be^{i\omega_2 t}, \qquad (2)$$

where $i\omega_1$, $i\omega_2$ are the roots of the characteristic polynomial $\lambda^2 - ib\lambda - a = 0$. These roots are purely imaginary if $b^2 > 4a$, which happens if the spin $L$ is large enough (here, $A$ and $B$ are complex numbers). If we put ourselves in the frame that is rotating with angular velocity $\omega_1$, then the motion (2) will acquire a clockwise rotation from our point of view:

$$e^{-i\omega_1 t} z = A + Be^{i(\omega_2 - \omega_1)t}.$$

This rotation is a uniform circular motion centered at $A$.

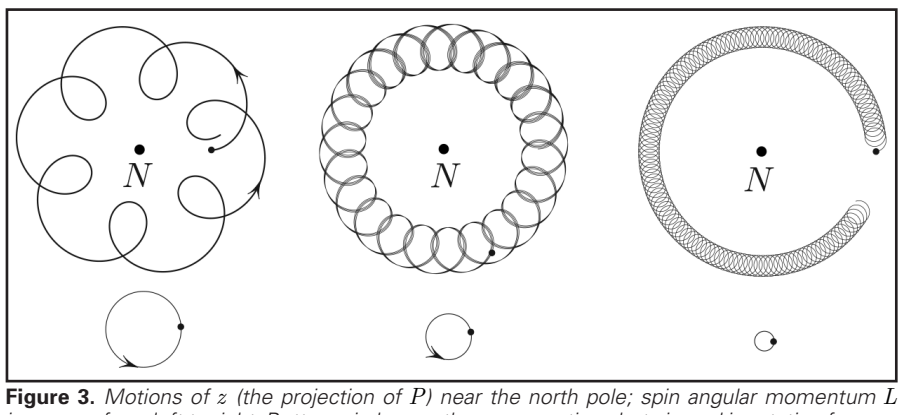



**Figure 3.** *Motions of $z$ (the projection of $P$) near the north pole; spin angular momentum $L$ increases from left to right. Bottom circles are the same motions but viewed in rotating frames.*

MATHEMATICAL CURIOSITIES
*By Mark Levi*

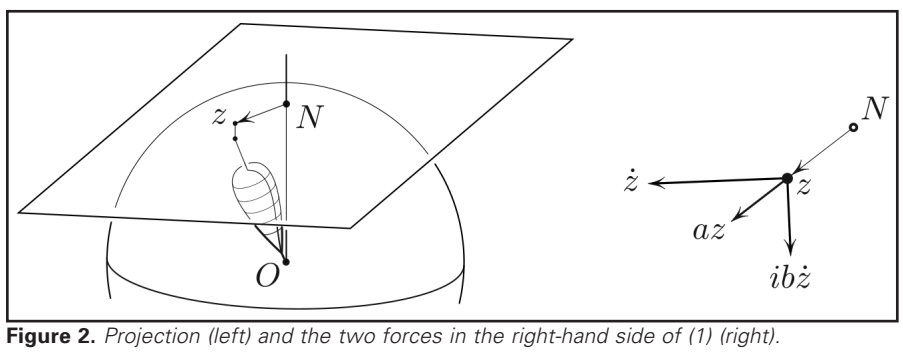



**Figure 2.** *Projection (left) and the two forces in the right-hand side of (1) (right).*

*The figures in this article were provided by the author.*

## References

[1] Arnold, V.I. (1989). *Mathematical methods of classical mechanics* (2nd ed). In *Graduate texts in mathematics*. New York, NY: Springer-Verlag.

[2] Cox, G., & Levi, M. (2016). Gaussian curvature and gyroscopes. *Comm. Pure Appl. Math.*, *71*(5), 938-952.

[3] Goldstein, H. (1980). *Classical mechanics*. Reading, MA: Addison-Wesley.

*Mark Levi (levi@math.psu.edu) is a professor of mathematics at the Pennsylvania State University.*

# The Rational Krylov Toolbox: Nonlinear Rational Approximation

*By Stefan Güttel*

In the April 2022 issue of *SIAM News*, I provided a brief introduction to rational Krylov methods—a critical tool in the field of scientific computing—and discussed several applications.[1] Now I focus on the RKFIT method for nonlinear rational approximation, one of the core algorithms in the freely available MATLAB Rational Krylov Toolbox (RKToolbox).[2] The toolbox comprises an extensive collection of examples that users can explore and modify; here they are printed in `typewriter font` for easy identification.

```
N = 100; A = gallery('grcar', N);
F = expm(A); b = ones(N, 1);
xi = inf(1, 4); % initial pole guess
[xi, ratfun] = rkfit(F, A, b, xi, 'real');
```

**Figure 1.** *Computation of a degree-4 RKFIT approximant to a matrix exponential.*

## Nonlinear Rational Approximation

The purpose of RKFIT is to solve nonlinear least-squares problems

$$\|Fb - r(A)b\|_2 \to \min_r, \qquad (1)$$

where $A$ and $F$ are given square matrices of the same size and $b$ is a given vector of compatible size. In the simplest case, the minimization in (1) occurs over all rational functions $r$ of degree $m$ (i.e., quotients $r = p/q$ of two polynomials of degree $m$). This is a nonconvex minimization problem that may be ill posed, meaning that we can generally only ask for a solution that makes the left side of (1) "small."

The formulation (1) contains several familiar special cases, including rational approximation on a discrete set when both $A$ and $F$ are diagonal matrices, or multipoint Padé approximation when $A$ has nontrivial Jordan blocks. In many applications, $F = f(A)$ is a complicated matrix function of $A$; we thus aim to compute $r$ without resorting to $A$'s eigenvalues or other spectral information. RKFIT achieves this objective by iteratively transforming the matrices in a rational Arnoldi decomposition and solving least-squares problems with orthonormal rational Krylov bases; more details are available elsewhere [1, 3].

SOFTWARE AND PROGRAMMING

The use of RKFIT in the RKToolbox is straightforward; Figure 1 presents a basic example. This code computes a degree-4 rational function $r$ such that $r(A)b \approx Fb$ for the nonnormal Grcar matrix $A$, which is a popular test matrix in numerical analysis. The computed function is represented by an RKFUN object called `ratfun` that users can incorporate in further computations. Figure 2 checks the relative error $\|Fb - r(A)b\|_2 / \|Fb\|_2$ and displays the four poles of $r$.

RKFIT has a number of optional functionalities and parameters that individuals can list by typing `help rkfit`. For instance, instead of providing `F` as a dense matrix like in our basic example, it is typically more efficient to provide a function handle for the computation of matrix-vector products; it is also possible to enforce stable poles (`param.stable`). The RKToolbox contains several RKFIT-related examples, two of which I will discuss in some detail.

## Exponential Integration

Linear initial value problems such as $u' + Au = 0$, $u(0) = b$—where $A$ is a large sparse matrix—arise in many applications. In the RKToolbox demonstration at `example_expint.hmtl`,[3] we are concerned with the efficient solution of this problem with uniform accuracy over a given time interval, say $[T_0, T_1]$. The solution, which is given as $u(t) = e^{-At}b$, can conveniently be approximated by a family of rational functions in partial fraction form:

$$r^{(t)}(A)b = \sum_{i=1}^{m} \alpha_i(t)(\xi_i I - A)^{-1} b \approx e^{-At}b.$$

*See* **Rational Krylov** *on page 12*

```
>> norm(F*b - ratfun(A,b))/norm(F*b)
ans =
   3.0754e-06

>> poles(ratfun)
ans =
   7.3887 - 1.7665i
   7.3887 + 1.7665i
   5.8385 - 5.4127i
   5.8385 + 5.4127i
```

**Figure 2.** *Exploring an RKFIT approximant.*

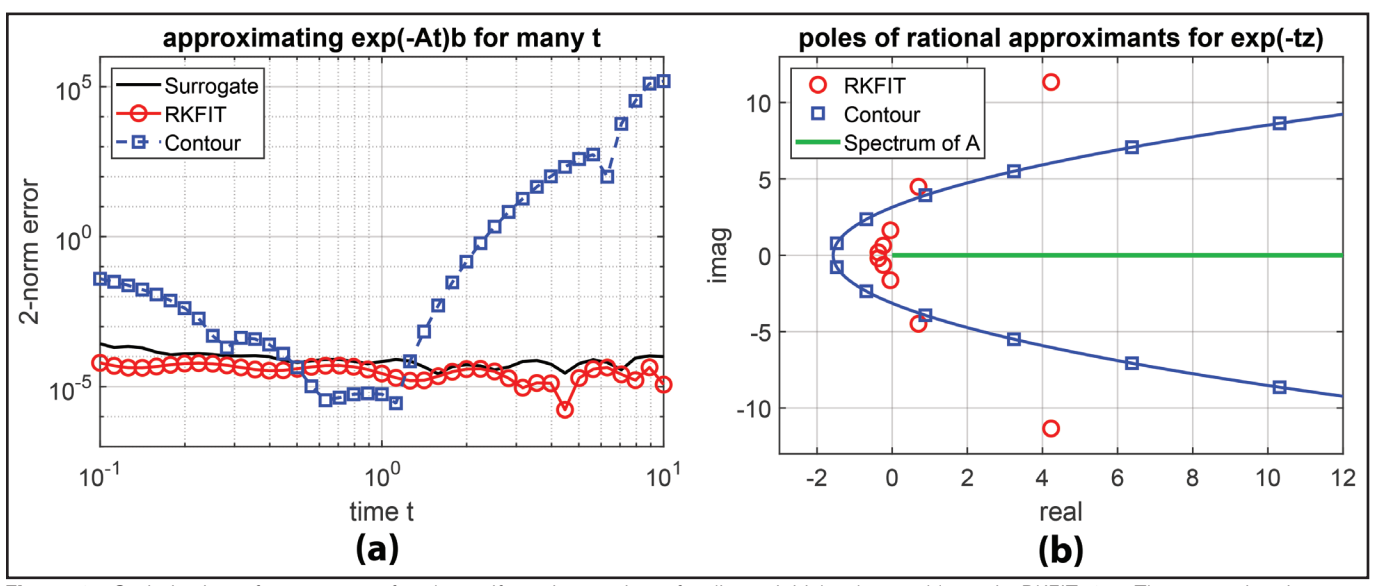



**Figure 3.** *Optimization of parameters for the uniform integration of a linear initial value problem via RKFIT.* **3a.** *The approximation error $\|e^{-At}b - r^{(t)}(A)b\|_2$ for a symmetric positive semidefinite matrix $A$ as a function of $t \in [0.1, 10]$ for RKFIT (represented by the red solid curve with circles) and a contour-based approach (represented by the blue dashed curve with squares). The solid black line depicts the errors of the RKFIT surrogate approximant.* **3b.** *The pole locations of the two families of rational approximants—RKFIT and contour-based—in the complex plane.*

[1] https://sinews.siam.org/Details-Page/the-rational-krylov-toolbox

[2] http://rktoolbox.org

[3] http://guettel.com/rktoolbox/examples/html/example_expint.html

# After Two Years, a Reunion in the Pacific Northwest

*By Thomas Humphries*

In addition to majestic forests, stunning mountains, and a beautiful coastline, the Pacific Northwest (PNW) region of North America is also home to a thriving applied mathematics community. The tradition of regular meetings between applied mathematicians in this part of the continent dates back at least 35 years, to the first Pacific Northwest Numerical Analysis Seminar[1] (PNWNAS) at the University of Washington in September 1987. This day-long conference, which consists of invited talks from speakers in both academia and industry, has convened almost every year since; the latest edition took place virtually in 2020.

In more recent years, collaboration between faculty at Oregon State University and Portland State University gave rise to the first CASCADE meeting in April 2014. This gathering—named for the mountainous Cascade Range in the PNW—uses a rodeo or circus format, in which the order and length of talks is determined on the day of the meeting based on which attendees volunteer to speak. The less-formal style makes it an ideal forum for students and early-career researchers to present their work. The meeting was later renamed CASCADE RAIN[2] (for Regional, Applied, Interdisciplinary, and Numerical mathematics) — an allusion to the PNW's infamously wet climate. The last iteration[3] occurred online at the start of the COVID-19 pandemic in 2020.

The PNW Section of SIAM[4] originated in 2016 to organize PNWNAS, CASCADE RAIN, and related events under one umbrella and make them accessible to students and participants beyond the major research universities in the region. The Section encompasses a large geographical area that includes the U.S. states of Washington, Oregon, and Idaho, along with the Canadian province of British Columbia. It organizes virtual seminars[5] by distinguished speakers several times a year to enable participation from a wide variety of members. Seminar topics have included tsunami modeling (see Figure 1), mathematical modeling of biological cells, clustering of complex networks, and ethical allocation of ventilators during COVID-19.

The PNW Section's most prominent event is its biennial meeting, the first of which was held in October 2017 at Oregon State University; the second conference took place at Seattle University two years later. Planning for the 3rd Biennial Meeting of the SIAM PNW Section[6] has been especially challenging in light of the ongoing COVID-19 pandemic. Although the gathering was originally scheduled for the fall of 2021, the organizers opted to postpone the meeting to the spring of 2022 due to continually high COVID-19 case counts. Subsequent logistical issues and scheduling conflicts resulted in several date changes and one venue change, but the organizing committee is now pleased to officially host the next biennial meeting from May 20-22 at Washington State University in Vancouver, Wash. Due to continued uncertainty surrounding the COVID-19 situation, the conference will employ a hybrid format; registered participants can choose to attend on campus or remotely via Zoom.

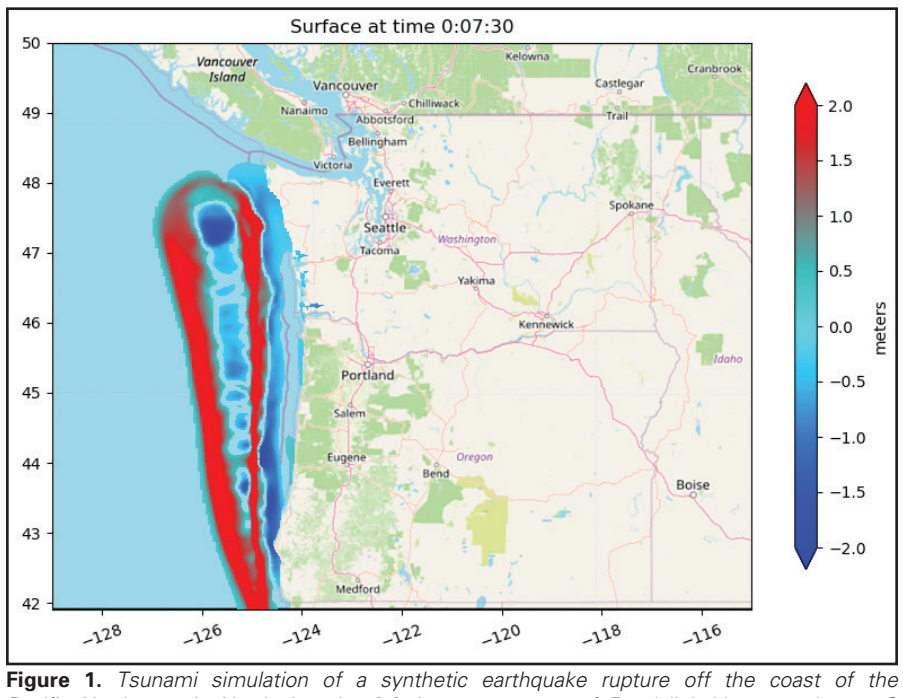



**Figure 1.** *Tsunami simulation of a synthetic earthquake rupture off the coast of the Pacific Northwest in North America [1]. Image courtesy of Randall LeVeque and map © OpenStreetMap contributors.*

Four plenary speakers—Tegan Emerson (Pacific Northwest National Laboratory), Jodi Mead (Boise State University), Jessica Stockdale (Simon Fraser University), and Jevin West (University of Washington)—will headline the meeting. They will present on topics such as COVID-19 modeling, the spread of viral misinformation, and topological data analysis. In addition, a number of minisymposia and contributed talks—organized by researchers from across the region—will address subjects like mathematical biology, computational fluid dynamics, numerical analysis, computational partial differential equations, and imaging science. Furthermore, a two-hour, in-person poster session on May 21 will allow students to practice their presentation skills and compete for prizes. We look forward to seeing many of our colleagues in person again after two long years of virtual meetings!

## References

[1] Melgar, D., LeVeque, R.J., Dreger, D.S., & Allen, R.M. (2016). Kinematic rupture scenarios and synthetic displacement data: An example application to the Cascadia subduction zone. *J. Geophys. Res. Solid Earth*, *121*(9), 6658-6674.

*Thomas Humphries is an assistant professor in the Division of Engineering and Mathematics in the School of Science, Technology, Engineering, and Mathematics at the University of Washington Bothell. He has served as secretary of the Pacific Northwest Section of SIAM since 2020.*

[1] https://sites.google.com/view/pnwnas

[2] https://sites.google.com/site/cascaderainmeetings/home

[3] https://sites.google.com/view/cascaderain2020

[4] https://sites.google.com/site/siampnwsection/home

[5] https://sites.google.com/site/siampnwsection/seminar-series

[6] https://sites.google.com/view/siampnw21

# Anticipating the 2023 SIAM Conference on Computational Science and Engineering in Amsterdam

*By Hans De Sterck, Karen Devine, Dirk Hartmann, Sarah Knepper, Wil Schilders, and Kees Vuik*

Goedendag! It's time to dust off your Dutch—or at least your passport—because the 2023 SIAM Conference on Computational Science and Engineering[1] (CSE23) is set to take place from February 26 to March 3, 2023, in Amsterdam, the Netherlands. CSE23 will be the first SIAM CSE conference outside of the U.S. in the meeting's 20-plus-year history; it will also be the first in-person iteration of CSE since CSE19[2] in Spokane, Wash. Stroll along the canals of Amsterdam while you catch up with colleagues, build your professional networks, and make new friends!

The biennial CSE meeting is historically SIAM's largest conference. It brings together applied mathematicians, computer scientists, domain scientists, and engineers to facilitate in-depth technical discussions about a wide variety of large-scale problems in computational science and engineering. The advanced models, simulations, and algorithms that CSE practitioners create inspire applications in fields that range from aerospace and chemistry to medicine, finance, and marketing. The broad audience that utilizes these applications thus fosters the interdisciplinary culture of CSE and helps train the next generation of scientists.

Some of the many themes of CSE23 include the following:

- Artificial intelligence and CSE
- High-performance computing
- Scalable linear and nonlinear solvers
- Reduced order modeling
- Data science for CSE
- Scientific machine learning
- Graphs and networks
- CSE in industry, including optics, energy, electronics, aerospace, and automobiles.

These and numerous other topics will manifest in the eight plenary talks that experts will present throughout the conference. Multiple minitutorials will provide opportunities for attendees to thoroughly explore subjects of interest in CSE, including software and tools. Engaging poster sessions will inspire individual and small group discussions that address state-of-the-art research. Finally, an awards ceremony with lectures from prize recipients will wrap up the week-long conference.

Other highlights at CSE23 include multiple panel discussions. In addition to the standard early- and mid-career panels (perennial favorites at the conference), another panel will focus on diversity — including the retention and long-term success of diverse employees as well as the integration of diversity into successful teams. In addition, a forward-looking panel composed of experts in the field will identify future trends in CSE.

Another first for CSE23 is the addition of a Hackathon that will take place alongside the conference. During this event, teams of students and their mentors will work on real-world industrial problems, meet other members of the future workforce, and learn about the complex problems that plague today's organizations.

While the plenary talks, panels, and minitutorials comprise the backbone of the conference, the bulk of the program consists of the multitude of minisymposia talks, contributed lectures, and posters that will be submitted and presented by the attendees themselves. We encourage *SIAM News* readers to submit their proposals once the call for participation is posted in May 2022.

Opportunities are available for institutions and companies to sponsor CSE23 activities and even host multiday booths at the conference. Sponsorship details and forms are accessible via the conference page.[3]

While SIAM membership[4] is not required for conference registration, it does provide a discount. An additional discount is available to members of the SIAM Activity Group on Computational Science and Engineering;[5] attendees can also save money by registering early.

CSE23 promises to be a great meeting that will unite CSE practitioners from all academic and career levels in a unique European venue. We hope to see you in Amsterdam, and tot dan!

*Hans De Sterck, Karen Devine, Dirk Hartmann, Sarah Knepper, Wil Schilders, and Kees Vuik are members of the Organizing Committee for the 2023 SIAM Conference on Computational Science and Engineering. Hans De Sterck is a professor of applied mathematics at the University of Waterloo and director of Waterloo's Centre for Computational Mathematics in Industry and Commerce. Karen Devine recently retired as a distinguished member of technical staff in the Center for Computing Research at Sandia National Laboratories. Dirk Hartmann is a Siemens Technical Fellow at Siemens Digital Industries Software. Sarah Knepper is an engineering manager at Intel Corporation. Wil Schilders is a professor of scientific computing for industry at Eindhoven University of Technology and executive director of the Dutch Platform for Mathematics. Kees Vuik is a professor of numerical analysis at the Delft University of Technology and scientific director of the Delft High Performance Computing Centre.*

[1] https://www.siam.org/conferences/cm/conference/cse23

[2] https://www.siam.org/conferences/cm/conference/cse19

[3] https://www.siam.org/conferences/cm/program/exhibits/cse23-exhibits

[4] https://go.siam.org/bVLwuH

[5] https://www.siam.org/membership/activity-groups/detail/computational-science-and-engineering


*The 2023 SIAM Conference on Computational Science and Engineering (CSE23) will take place from February 26 to March 3, 2023, in Amsterdam, the Netherlands.*

## Rational Krylov

*Continued from page 10*

Note that the *poles* $\xi_i$ of $r^{(t)}$ are independent of $t$ and the evaluation of $r^{(t)}(A)b$ hence requires the solution of $m$ shifted linear systems $(\xi_i I - A)x_i = b$, which are independent of the number of time points $t$ for which we want to evaluate. Furthermore, all of these linear systems are decoupled and can be solved in parallel.

A popular approach for obtaining families of such rational approximants is to apply a quadrature discretization to the Cauchy integral formula

$$e^{-At} = \frac{1}{2\pi \mathrm{i}} \int_\Gamma e^{-\zeta t}(\zeta I - A)^{-1} \mathrm{d}\zeta$$

with an appropriately chosen contour $\Gamma$ that encloses $A$'s eigenvalues [4]. Alternatively, we can run RKFIT to numerically compute a family of rational approximants so that the error $\| e^{-At}b - r^{(t)}(A)b \|_2$ is uniformly small for all $t \in [T_0, T_1]$. To that end, RKFIT employs a *surrogate approach* wherein $A$ and $b$ are replaced by a "simpler" matrix $\hat{A}$ and vector $\hat{b}$. The surrogate matrix $\hat{A}$ should capture $A$'s spectral properties while being easy to compute with. For instance, one could obtain $\hat{A}$ from a coarser discretization if $A$ arises from the spatial discretization of a differential operator. Figure 3 (on page 10) illustrates an example of this concept [1]. Here, the matrix $A$ is symmetric positive semidefinite and the time interval of interest is $[T_0, T_1] = [0.1, 10]$.

Figure 3a displays the approximation error of both the original and surrogate problems. The RKFIT family of rational approximants $r^{(t)}(A)b$ achieves a near-uniform error over the entire time interval. We also compare this approach to a contour integration-based technique that is optimized for a single time parameter $t = 1$, where the poles of the rational function lie on a prescribed contour (see Figure 3b, on page 10). The contour approach achieves a high accuracy at $t = 1$, but the accuracy deteriorates significantly at other time points.

### Compression of Layered Waveguides

The example at `example_ehcompress.html`[4] relates to absorbing boundary conditions for wave problems and the optimization of transmission conditions in domain decomposition methods. Consider a two-dimensional waveguide with varying wave number in the horizontal direction (see Figure 4a). Depending on the structure of the layers, the Dirichlet-to-Neumann (DtN) map of this waveguide, $f(A)$, may be a highly irregular matrix function where $f$ has many singularities (so-called scattering poles) near $A$'s eigenvalues. It is therefore impossible to construct a uniform rational approximant $r(A) \approx f(A)$ on $A$'s spectral region. We recently proposed the computation of a low-order RKFIT approximant $r(A)b \approx f(A)b$ with a random probing vector $b$ [2]. The continued fraction form of $r$ is equivalent to a finite-difference representation of the DtN map; this representation can be appended to an existing discretization as an artificial boundary condition.

An important property of RKFIT approximants is their inherent spectral adaptation. Even though one does not explicitly need $A$'s eigenvalues to run RKFIT, the computed approximant will resolve the function $f$ more accurately in some parts of the spectrum than in others (see Figure 4b). This spectral adaptation effect allows for the construction of finite-difference grids that undercut the Nyquist limit for discretizations of wave problems [2].

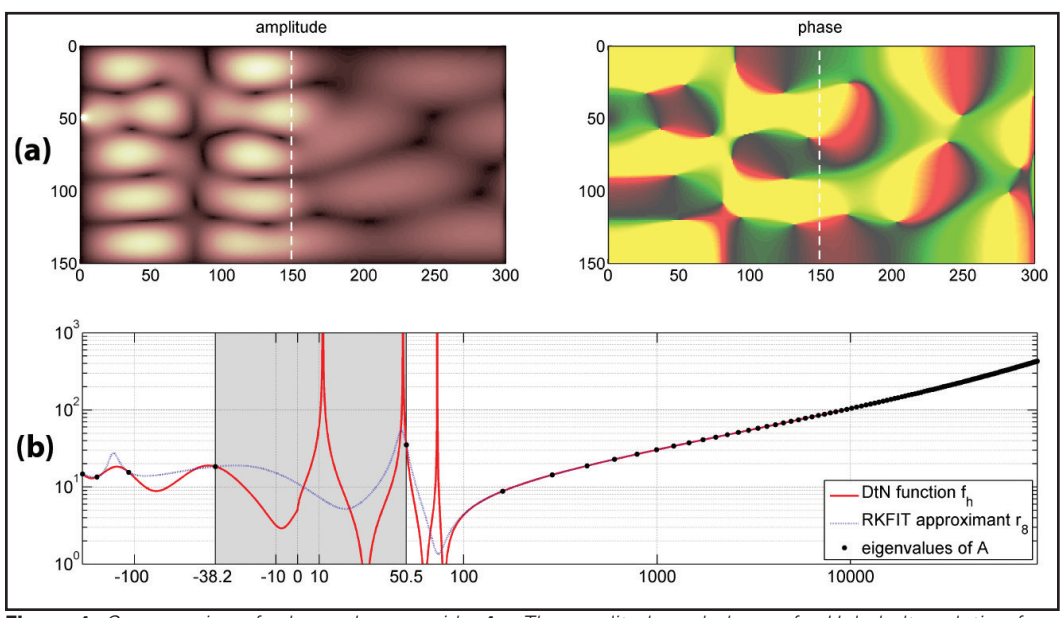



**Figure 4.** *Compression of a layered waveguide.* **4a.** *The amplitude and phase of a Helmholtz solution for a two-dimensional waveguide with varying coefficients (wave numbers) in the horizontal $x$-direction. The coefficient jump occurs at $x = 150$, as indicated by the vertical dashed line.* **4b.** *A plot of the Dirichlet-to-Neumann (DtN) function $f$ of the waveguide (represented by the solid red line) over the spectral interval of an indefinite matrix $A$ (the trace operator at the $x = 0$ interface). The RKFIT approximant of degree $m = 8$ (represented by the dotted blue curve) is not uniformly close to $f$ on the spectral interval of $A$. It exhibits spectral adaptation to some of $A$'s eigenvalues (represented by the black dots).*

### Summary

Here I have presented a brief introduction to the RKFIT method for nonlinear rational approximation and discussed two sample applications. The RKToolbox provides some additional RKFIT examples that are related to model order reduction (`example_frequency.html`[5] and `example_iss.html`[6]) and graph label propagation (`example_digits.html`[7]).

*All figures are courtesy of the author.*

### References

[1] Berljafa, M., & Güttel, S. (2017). The RKFIT algorithm for nonlinear rational approximation. *SIAM J. Sci. Comput.*, *39*(5), A2049-A2071.

[2] Druskin, V., Güttel, S., & Knizhnerman, L. (2022). Model order reduction of layered waveguides via rational Krylov fitting. *BIT Numer. Math*. To appear.

[3] Güttel, S. (2022, April 1). The rational Krylov toolbox. *SIAM News*, *55*(3), p. 10.

[4] Trefethen, L.N., Weideman, J.A.C., & Schmelzer, T. (2006). Talbot quadratures and rational approximations. *BIT Numer. Math.*, *46*(3), 653-670.

*Stefan Güttel is a professor of applied mathematics at the University of Manchester and a Fellow of the Alan Turing Institute. He received SIAM's 2021 James H. Wilkinson Prize in Numerical Analysis and Scientific Computing.*

[4] http://guettel.com/rktoolbox/examples/html/example_ehcompress.html

[5] http://guettel.com/rktoolbox/examples/html/example_frequency.html

[6] http://guettel.com/rktoolbox/examples/html/example_iss.html

[7] http://guettel.com/rktoolbox/examples/html/example_digits.html