## Research Overview for Christopher J. Hillar

I work on a wide range of problems that arise from other areas of mathematics and the physical sciences. Currently, I am focused on using mathematical and computational tools to solve basic problems in theoretical neuroscience, and in this regard, I have begun collaborations with scientists at the Redwood Center for Theoretical Neuroscience and mathematicians at U.C. Berkeley. I am also interested in theoretical questions involving semidefinite programming, optimization, and computational algebra.

The following is a description of several interrelated lines of research in which I will actively participate in the coming years. The first three sections contain very brief discussions of topics related to theoretical neuroscience that I have only begun exploring in recent months. The final sections describe more theoretical studies that I have been investigating in recent years and therefore contain more detailed descriptions.

### 1. Sparse coding and compressed sensing

Sparse coding refers to the process of representing a real vector input (such as an image) as a sparse linear combination of an overcomplete set of vectors (called a *sparse basis*). Here, overcomplete refers to the fact that there are many more vectors in the sparse basis than the vector space dimension. In recent years, it has been discovered that sparse coding could be a model for how the brain's visual cortex represents low-level features in images (such as edges) as the receptive fields of neurons [56].

Also, in recent years, there has been a flurry of activity surrounding what is called compressed sensing (see the references in [9]). The idea is that by taking a small number of random projections (into a much smaller space) of an input vector, one retains enough information to reconstruct accurately the input by performing an (efficient) $\ell_2$ optimization with a sparse $\ell_1$ penalty.

With Fritz Sommer (Redwood Center for Theoretical Neuroscience) and his student Will Coulter, we are trying to combine these two techniques for representing data. The motivation from neuroscience is that overcompleteness and compression might be playing dual roles between distant regions in the brain. On the one hand, a sparse basis is a way to explain input from a small selection of a large number of (possible) causes, while on the other, data is transmitted more efficiently by taking a small number of random projections into a space of size commensurate to the (presumably small) number of causes actively making up the input. The mathematics of the investigation involves techniques from optimization theory, matrix analysis, and (potentially) semidefinite programming [65].

### 2. Phase modeling

Circular variables such as phase or orientation have been used effectively for representing complex physical phenomenon and in the analysis and processing of signals. Countless physical systems are effectively represented using phase variables. Coupled oscillator systems are prevalent in classical physics as a canonical model of systems ranging from coupled pendula to coupled Josephson junctions. Oscillator models have also been effective at describing coupled behavior in nature: chemical reaction diffusion systems, heart-lung and circadian rhythms, and even the coupling of firefly luminescence can all be described with phase variables. In engineering, phase has played a key role in signal representation. From classical Fourrier analysis to modern techniques in image representation (for example [19, 62, 25]), phase provides a useful representation.

Within neuroscience oscillatory dynamics and phase variables have had an especially interesting history. Oscillatory dynamics played a central role in many early theories of large-scale brain dynamics [26], and oscillatory dynamics have recently received widespread interest [27, 69, 14]. Network oscillations are hypothesized to be functionally involved in a wide range of tasks, such as representation of sensory information, regulating the flow of information, learning and memory-recall of information, and *binding* of distributed information. Clearly, phase is of central importance to the field.

Motivated by observations of empirical data, Cadieu and Koepsell (Redwood Center for Theoretical Neuroscience) have introduced the following $d$-dimensional multivariate phase distribution:

$$p(\boldsymbol{\theta}|\mathbf{K}) = \frac{1}{Z(\mathbf{K})} \exp[-E(\boldsymbol{\theta};\mathbf{K})] = \frac{1}{Z(\mathbf{K})} \exp\left[-\tfrac{1}{2}\,\mathbf{x}^*\mathbf{K}\mathbf{x}\right], \tag{2.1}$$

where $\mathbf{x}$ is the $d$-dimensional complex vector with components $x_i = e^{j\theta_i}$, $\mathbf{K}$ is a $d \times d$-dimensional Hermitian matrix, and $Z(\mathbf{K})$ is the normalization constant needed to assure that the probability integrates to one. Note that it is non-trivial to determine the normalization $Z(\mathbf{K})$. With Charles Cadieu and Kilian Koepsell, I shall study the mathematics of this distribution; in particular, we shall develop algorithms for determining the parameters in $\mathbf{K}$ given data sampled from the model.

## 3. Computational complexity of tensor problems

There has been much work recently on "tensor methods" in computer vision, data analysis, machine learning, scientific computing, neuroscience, and other areas. The idea of using tensors for numerical computing is attractive. Nearly all problems in computational science and engineering may eventually be reduced to one or more standard problems involving matrices: systems of linear equations, least squares problems, eigenvalue problems, low-rank approximations, matrix decompositions such as LU, QR, EVD, SVD, etc. If similar problems for tensors of higher order may be solved effectively, then one would have substantially enlarged the arsenal of fundamental tools in numerical computations.

With Lek-Heng Lim (U.C. Berkeley), we are investigating the computational complexity of tensor analogues of many problems that are readily computable in numerical linear algebra. Johan Håstad has already shown that tensor rank is NP-hard over $\mathbb{Q}$ and NP-complete over finite fields [32]. We are extending the list to include many more multilinear generalizations of standard problems in numerical linear algebra. Furthermore, these multilinear problems are not just natural extensions of their linear cousins but have all surfaced in recent applications.

Let $\mathbb{F}$ be the field $\mathbb{R}$ or $\mathbb{C}$. Also, let $\mathcal{A} \in \mathbb{F}^{l\times m\times n}$ be a 3-tensor and let $\mathcal{S} \in \mathbb{F}^{n\times n\times n}$ be a symmetric 3-tensor. The list of problems we shall analyze includes the following:

- Computing the spectral norm of a 3-tensor:

$$\sup_{\mathbf{x},\mathbf{y},\mathbf{z}\neq\mathbf{0}} \frac{|\mathcal{A}(\mathbf{x},\mathbf{y},\mathbf{z})|}{\|\mathbf{x}\|_2\|\mathbf{y}\|_2\|\mathbf{z}\|_2}.$$

- Computing the spectral norm of a symmeric 3-tensor:
$$\sup_{\mathbf{x}\neq\mathbf{0}} \frac{|\mathcal{S}(\mathbf{x},\mathbf{x},\mathbf{x})|}{\|\mathbf{x}\|_2^3}.$$
- Computing a best rank-1 approximation to a tensor:
$$\min_{\mathbf{x},\mathbf{y},\mathbf{z}} \|\mathcal{A} - \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_F.$$
- Computing a best rank-1 approximation to a symmetric tensor:
$$\min_{\mathbf{x}} \|\mathcal{S} - \mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\|_F.$$
- Computing a singular value, or, given a singular value, compute the corresponding singular vectors of a 3-tensor. These are defined as stationary values and stationary points of the trilinear Rayleigh quotient
$$\frac{\mathcal{A}(\mathbf{x},\mathbf{y},\mathbf{z})}{\|\mathbf{x}\|_p\|\mathbf{y}\|_p\|\mathbf{z}\|_p},$$
where $p = 2$ or 3.
- Computing an eigenvalue, or, given an eigenvalue, compute a corresponding eigenvector of a symmetric 3-tensor. These are defined as stationary values and stationary points of the trilinear Rayleigh quotient
$$\frac{\mathcal{S}(\mathbf{x},\mathbf{x},\mathbf{x})}{\|\mathbf{x}\|_p^p},$$
where $p = 2$ or 3.
- Solving a system of bilinear equations in the exact sense
$$\mathcal{A}(\mathbf{x},\mathbf{y},I) = \mathbf{b}$$
or approximating it in the least-squares sense
$$\min_{\mathbf{x},\mathbf{y}} \|\mathcal{A}(\mathbf{x},\mathbf{y},I) - \mathbf{b}\|_2.$$

We have several results that indicate most, if not all, of these are NP hard – both in the traditional Cook-Karp-Levin sense [16, 41, 49] and also the Blum-Shub-Smale sense [8]. The main tools used in our analysis are algebraic formulations of combinatorial problems such as 3-colorability in complexity theory [5]. I have previous experience in this regard; for instance, with Windfeldt, we showed [39] that there is an algebraic characterization of unique $k$-colorability of a graph, and we used this result to verify a counterexample of Akbari, Mirrokni, and Sadjad [1] to a conjecture of Xu [71]. See Section 6 below for more details.

## 4. Real Algebraic Geometry and Optimization

4.1. **Rational sums of squares.** In recent years, techniques from semidefinite programming have produced numerical algorithms for expressing positive semidefinite polynomials as sums of squares. These algorithms have many applications in optimization, control theory, quadratic programming, and matrix analysis [57, 58, 59, 60, 61]. Moreover, such representations aid in the computation of the real locus of a polynomial. For a noncommutative application of these techniques to a famous trace conjecture, see the next section, which discusses the papers [12, 29, 34, 42, 47, 68].

One major drawback with these algorithms is that their output is, in general, numerical. For many applications, however, exact polynomial identities are needed. In this regard, Sturmfels has asked the following question.

**Question 4.1** (Sturmfels)**.** *If $f \in \mathbb{Q}[x_1, \ldots, x_n]$ is a sum of squares in $\mathbb{R}[x_1, \ldots, x_n]$, then is $f$ also a sum of squares in $\mathbb{Q}[x_1, \ldots, x_n]$?*

This question has a positive answer in the univariate case due to results of Landau [46] and Pourchet [63]; an algorithmic solution is due to Schweighofer [68]. It follows from a famous theorem of Artin [64] that if $f \in \mathbb{Q}[x_1, \ldots, x_n]$ is a sum of squares of rational functions in $\mathbb{R}(x_1, \ldots, x_n)$, then it is a sum of squares in $\mathbb{Q}(x_1, \ldots, x_n)$. Moreover, from the work of Voevodsky on the Milnor conjectures, it is known that $2^{n+2}$ such squares suffice [45]. However, the transition from rational functions to polynomials is often a very delicate one. For instance, not every polynomial that is a sum of squares of rational functions is a sum of squares of polynomials [64].

More generally, Sturmfels is interested in the algebraic degree [55] of maximizing a linear functional over the space of all sum of squares representations of a given polynomial that is a sum of squares. In the special case of Question 4.1, a positive answer signifies an algebraic degree of 1 for this optimization problem.

General theory reduces Question 4.1 to one involving real algebraic numbers. Recently, I made progress in the multivariate case when the coefficients lie in a totally real number field $K$. My main theorem in [33] is the following.

**Theorem 4.2.** *Let $K$ be a totally real number field with Galois closure $L$ and let $R$ be a commutative $\mathbb{Q}$-algebra. If $f \in R$ is a sum of $m$ squares in $R \otimes_{\mathbb{Q}} K$, then $f$ is a sum of $4m \cdot 2^{[L:\mathbb{Q}]}\binom{[L:\mathbb{Q}]+1}{2}$ squares in $R$.*

My proof of Theorem 4.2 was constructive. It is known [15] that arbitrarily large numbers of squares are necessary to represent any sum of squares over $\mathbb{R}[x_1, \ldots, x_n]$, $n > 1$, making a fixed bound (for a given $n$) as in the rational function case impossible.

I am working on the generalization of this result to any real algebraic extension of $\mathbb{Q}$ in order to give a complete answer to Question 4.1. This work will settle the important question of how much limitation one has in using semidefinite techniques for finding algebraic certificates of nonnegativity.

4.2. **The BMV trace conjecture.** In 1975, while studying partition functions of quantum mechanical systems, Bessis, Moussa, and Villania formulated a conjecture regarding a positivity property of traces of matrices [7]. If this property holds, explicit error bounds in a sequence of Padé approximants follow. Let $A$ and $B$ be $n \times n$ Hermitian matrices with $B$ positive semidefinite, and let

$$\phi^{A,B}(t) = \mathrm{Tr}[\exp{(A - tB)}].$$

The original formulation of the conjecture asserts that $\phi^{A,B}$ is completely monotone.

Since the conjecture was introduced in [7], many partial results and extensive computational experimentation have been given all in favor of the conjecture's validity. For instance, there have been variational approaches [17, 18, 34, 35], hypergeometric approaches [23], free probability approaches [24], techniques using matrix analysis [30, 40, 52], and most recently noncommutative sum of squares approaches [12, 29, 43, 47] that have been combined with the technology of semidefinite programming [42]. However, despite much work, very little is known about the problem, and it has remained unresolved except in very special cases. Recently, Lieb and Seiringer in [48], and as previously communicated to us [40], have reformulated the conjecture of [7] as a question about the traces of certain sums of words in two positive definite matrices.

**Conjecture 4.3** (BMV). *The polynomial $p(t) = \mathrm{Tr}\,[(A+tB)^m]$ has all nonnegative coefficients whenever $A$ and $B$ are $n \times n$ positive semidefinite matrices.*

The coefficient of $t^k$ in $p(t)$ is the trace of the sum, $S_{m,k}(A,B)$, of all words of length $m$ in $A$ and $B$, in which $k$ $B$'s appear. In [40], among other things, it was noted that, for $m < 6$, each constituent word in $S_{m,k}(A,B)$ has nonnegative trace. Thus, the above conjecture is valid for $m < 6$ and arbitrary positive integers $n$. It was also noted in [40] (see also [7]) that the conjecture is valid for arbitrary $m$ and $n < 3$. Thus, the first case in which prior methods did not apply and the conjecture was in doubt, is $m = 6$ and $n = 3$. Even in this case, all coefficients, except $\mathrm{Tr}[S_{6,3}(A,B)]$, were known to be nonnegative (also as shown in [40]). It was only recently [35], using heavy computation, that Johnson and I showed this remaining coefficient to be nonnegative.

Much of the subtlety of Conjecture 4.3 lies in the fact that $S_{m,k}(A,B)$ need not have all nonnegative eigenvalues, and in addition that some terms within the sum defining $S_{m,k}(A,B)$ can have negative trace. This later fact was only proved recently in work with Johnson [40], in which we disproved the conjecture [40] that all positive definite words in two letters have positive trace.

In [34], I made progress on the conjecture with the following theorem.

**Theorem 4.4.** *Suppose that there exist integers $m', k'$ and $n \times n$ positive definite matrices $A$ and $B$ such that $\mathrm{Tr}[S_{m',k'}(A,B)] < 0$. Then, for any $m \geq m'$ and $k \geq k'$ such that $m - k \geq m' - k'$, there exist $n \times n$ positive definite $A$ and $B$ making $\mathrm{Tr}[S_{m,k}(A,B)]$ negative.*

**Corollary 4.5.** *If the Bessis-Moussa-Villani conjecture is true for some exponent $m_0$, then it is also true for all $m < m_0$.*

Corollary 4.5 motivates a general program to solve the BMV conjecture, and there is evidence that this approach is more than a theoretical possibility. For instance, Hägele [29] has used this approach and Corollary 4.5 to prove the conjecture for all $m \leq 7$ (and all $n$). Inspired by Hägele's ideas, Klep and Schweighofer [42] used semidefinite programming and noncommutative sums of squares techniques to prove the conjecture for all $m \leq 13$. One of their motivations was the Connes' embedding conjecture on von Neumann algebras [43]. It should be noted that these techniques provably fail [29] for the difficult $m = 6$ case, making the appeal to Corollary 4.5 fundamental. Other work along these lines appears in the papers of Burgdorf [12] and Landweber-Speer [47].

Another approach is to use the following theorem found in my paper [34]. It characterizes the BMV conjecture in terms of the eigenvalues of the matrix $S_{m,k}(A,B)$ and resembles the Perron-Frobenius theorem for nonnegative matrices.

**Theorem 4.6.** *Fix positive integers $m \geq k$ and $n$. Then, $\mathrm{Tr}[S_{m,k}(A,B)] \geq 0$ for all positive semidefinite $A$,$B$ if and only if for all positive semidefinite $A$,$B$, the matrix $S_{m,k}(A,B)$ either has a positive eigenvalue or is the zero matrix.*

It follows that to prove the BMV conjecture, it is enough to show the positivity of only one of the eigenvalues of $S_{m,k}(A, B)$, rather than the sum of all of them. I propose to continue the study of this conjecture in light of these new methods and approaches.

## 5. Finiteness questions in rings with infinite Krull dimension

5.1. **Ideals of Algebraic Relations.** In chemistry [51, 66, 67] and algebraic statistics [22], a motivating problem is to determine the algebraic relations between experimental measurements. In this regard, Sturmfels has asked whether, up to symmetry, there are finitely many of them that generate the others. We discuss the mathematics of this problem and an approach by Aschenbrenner and myself for solving it.

Fix a natural number $k \geq 1$. Given a positive integer $n$, we denote by $\langle n \rangle^k$ the set of all ordered $k$-element subsets of $\{1, \dots, n\}$. Let $K$ be a field, and for $n \geq k$ consider the polynomial ring $R_n = K\big[\{x_{\boldsymbol{u}}\}_{\boldsymbol{u} \in \langle n \rangle^k}\big]$. We let $\mathfrak{S}_n$ act on $\langle n \rangle^k$ by

$$\sigma(u_1, \dots, u_k) = \big(\sigma(u_1), \dots, \sigma(u_k)\big).$$

This induces an action $(\sigma, x_{\boldsymbol{u}}) \mapsto \sigma x_{\boldsymbol{u}} = x_{\sigma \boldsymbol{u}}$ of $\mathfrak{S}_n$ on the indeterminates $x_{\boldsymbol{u}}$, which we extend to an action of $\mathfrak{S}_n$ on $R_n$ in the natural way. Set $R = \bigcup_{n \geq k} R_n$. Note that $R = K\big[\{x_{\boldsymbol{u}}\}_{\boldsymbol{u} \in \langle \Omega \rangle^k}\big]$, where $\Omega = \{1, 2, 3, \dots\}$ is the set of positive integers, and that the actions of $\mathfrak{S}_n$ on $R_n$ combine uniquely to an action of $\mathfrak{S}_\infty$ on $R$. Now let $f(y_1, \dots, y_k) \in K[y_1, \dots, y_k]$, let $t_1, t_2, \dots$ be an infinite sequence of pairwise distinct indeterminates over $K$, and for $n \geq k$ consider the $K$-algebra homomorphism

$$\phi_n \colon R_n \to K[t_1, \dots, t_n], \qquad x_{(u_1, \dots, u_k)} \mapsto f(t_{u_1}, \dots, t_{u_k}).$$

The ideal $Q_n = \ker \phi_n$ of $R_n$ determined by such a map is the prime ideal of algebraic relations between the quantities $f(t_{u_1}, \dots, t_{u_k})$. An important open problem is to understand the limiting behavior of such relations.

The ideals $Q_n$ form an increasing *chain* $Q_\circ : Q_k \subseteq Q_{k+1} \subseteq \cdots \subseteq Q_n \subseteq \cdots$. Such chains fail to stabilize in the usual sense; however, it is possible for them to stabilize "up to the action of the symmetric group", a concept we make precise below. Notice first that the chains induced by a polynomial $f$ are *invariant* under the action of the symmetric group in the sense that

$$\langle \mathfrak{S}_m Q_n \rangle \subseteq Q_m \text{ and } R_n \cap Q_m \subseteq Q_n \text{ for all } n \leq m.$$

Equivalently, the ideal $Q = \bigcup_{n \geq k} Q_n \subseteq R$ is invariant under the action of $\mathfrak{S}_\infty$. The stabilization definition alluded to above is as follows.

**Definition 5.1.** *A chain $Q_\circ$ stabilizes if there exists a positive integer $N$ such that*

$$\langle \mathfrak{S}_m Q_n \rangle = Q_m \qquad \textit{for all } m \geq n > N.$$

To put it another way, accounting for the natural action of the symmetric group, the ideals $Q_n$ are the same for large enough $n$. In applications, this would imply that there are only a finite number of "test relations" to check whether a series of measurements satisfies a hypothetical underlying model.

When $k = 1$, Aschenbrenner and I have shown that $R$ is Noetherian as an $R[\mathfrak{S}_\infty]$-module [4], which implies that any invariant chain stabilizes. When $k > 1$, however, $R$ is no longer Noetherian, making Sturmfels' question about stability much more subtle. In [4], we were able to prove a special case.

**Theorem 5.2.** *The sequence of kernels $Q_n$ induced by a square-free monomial $f \in K[y_1, \ldots, y_k]$ stabilizes. Moreover, a bound for when stabilization occurs is $N = 4k$.*

The proof of this result used in a special way the toric geometry that underlies this question. This theorem provides evidence for the following conjecture.

**Conjecture 5.3.** *The sequence of kernels induced by any monomial $f \in K[y_1, \ldots, y_k]$ stabilizes.*

We propose to settle this conjecture. A key step will be to generalize a theorem of Camina and Evans [13]. Namely, we will give a characterization of all the $\mathfrak{S}_\infty$-submodules of $\mathbb{Q}\langle\Omega\rangle^k$. We will then use this description to get precise information on the union $Q$ of toric ideals $Q_n$. These results will also be of independent interest.

5.2. **Symbolic Computation of Symmetric Gröbner Bases.** In computational algebra, one encounters the following general problem.

**Problem 5.4.** *Let $I$ be an ideal of a ring $R$ and let $f \in R$. Determine whether $f \in I$.*

When $R = K[x_1, \ldots, x_n]$ is a polynomial ring in $n$ indeterminates over a field $K$, this problem has a spectacular solution due to Buchberger [11].

**Theorem 5.5** (Buchberger)**.** *Let $I = \langle f_1, \ldots, f_m\rangle_R$ be an ideal of $R = K[x_1, \ldots, x_n]$. Then, there is a computable, finite set of polynomials $G$ such that for every polynomial $f$, we have $f \in I$ if and only if the polynomial reduction of $f$ with $G$ is* 0*.*

One remarkable feature of this result is that once such a *Gröbner basis* $G$ for $I$ is found, any new instance of the question "Is $f \in I$"? can be solved very quickly. Theorem 5.5 forms the backbone of the field of computational algebraic geometry.

We study a different but related membership problem Let $X = \{x_1, x_2, \ldots\}$ be an infinite collection of indeterminates, indexed by the positive integers, and let $\mathfrak{S}_\infty$ be the group of permutations of $X$. For a positive integer $N$, we will also let $\mathfrak{S}_N$ denote the set of permutations of $\{1, \ldots, N\}$. Fix a field $K$ and let $R = K[X]$ be the polynomial ring in the indeterminates $X$. The group $\mathfrak{S}_\infty$ acts naturally on $R$: if $\sigma \in \mathfrak{S}_\infty$ and $f \in K[x_1, \ldots, x_n]$,

$$\sigma f(x_1, \ldots, x_n) = f(x_{\sigma 1}, \ldots, x_{\sigma n}) \in R. \tag{5.1}$$

We motivate our discussion with the following concrete problem. Questions of this nature arise in applications to chemistry [51, 66, 67] and algebraic statistics [22].

**Problem 5.6.** *Let $f_1 = x_1^3 x_3 + x_1^2 x_2^3$ and $f_2 = x_2^2 x_3^2 - x_2^2 x_1 + x_1 x_3^2$ and consider the ideal $I = \langle \mathfrak{S}_\infty f_1, \mathfrak{S}_\infty f_2\rangle_R$ of $R = K[X]$ generated by all permutations of $f_1$ and $f_2$. Is the following polynomial involving* 10 *indeterminates in $I$?*

$$\begin{aligned} f \;&= -x_{10}^2 x_9^2 x_5^6 - 2x_{10}^2 x_9 x_8^3 x_5^5 - x_{10}^2 x_8^6 x_5^4 + 3x_{10}^2 x_8^2 + 3x_{10}^2 x_7 + 3x_{10} x_9 x_7 x_4^3 x_3^2 x_2^2 x_1 \\ &+ 3x_{10} x_9 x_7 x_4^3 x_3^2 x_1^2 - 3x_{10} x_9 x_7 x_4^3 x_2^2 x_1^2 - x_9^2 x_8^7 x_7 x_6 x_5^6 - 2x_9 x_8^{10} x_7 x_6 x_5^5 \\ &+ x_9 x_5^3 x_3 x_2 x_1^3 + x_9 x_5^3 x_2^4 x_1^2 + x_9 x_3 x_2^3 x_1^4 + x_9 x_2^6 x_1^3 - x_8^1 3 x_7 x_6 x_5^4 - 3x_8^2 x_7 \\ &+ x_7^2 x_6 x_3^3 x_2^7 + x_7^2 x_6 x_3^3 x_2^5 x_1 - x_7^2 x_6 x_3 x_2^7 x_1 + x_5 x_4^2 - 3x_5 x_3^2 + 2x_5 x_1^2 + x_4^2 x_3^2 \\ &- 2x_3^2 x_1^2 + 5x_3 x_1^5 + 5x_2^3 x_1^4. \end{aligned}$$

Naively, one could solve this problem using Buchberger's algorithm with truncated polynomial rings $R_n = K[x_1, \ldots, x_n]$. Namely, for each $n \geq 10$, compute a Gröbner basis $G_n$ for the ideal $I_n = \langle \mathfrak{S}_n f_1, \mathfrak{S}_n f_2 \rangle_{R_n}$, and reduce $f$ by $G_n$. There are several problems with this approach. For one, this method requires computation of many Gröbner bases (the bottleneck in any symbolic computation), the number of which depends on the number of indeterminates appearing in $f$. Additionally, it lacks the ability to solve new membership problems quickly, a powerful feature of Buchberger's technique.

Building on our work in [4], Aschenbrenner and I have been developing an algorithm that solves the general membership problem for symmetric ideals (such as those appearing in Problem 5.6) and has all of the important features of Buchberger's method. It is the first algorithm of its kind that we are aware of. We develop some notation.

Let $R[\mathfrak{S}_\infty]$ denote the (left) group ring of $\mathfrak{S}_\infty$ over $R$ with multiplication given by $f\sigma \cdot g\tau = fg(\sigma\tau)$ for $f, g \in R$ and $\sigma, \tau \in \mathfrak{S}_\infty$, and extended by linearity. The action (5.1) naturally gives $R$ the structure of a (left) module over the ring $R[\mathfrak{S}_\infty]$. An ideal $I \subseteq R$ is called *symmetric* if $\mathfrak{S}_\infty I := \{\sigma f : \sigma \in \mathfrak{S}_\infty,\ f \in I\} \subseteq I$. Symmetric ideals are then simply the $R[\mathfrak{S}_\infty]$-submodules of $R$.

We will also use the following notation. Let $B$ be a ring and let $G$ be a subset of a $B$-module $M$. Then $\langle f : f \in G \rangle_B$ will denote the $B$-submodule of $M$ generated by the elements of $G$. For instance, the invariant ideal $I = \langle x_1, x_2, \ldots \rangle_R$, as a module over the group ring $R[\mathfrak{S}_\infty]$, has the compact presentation $I = \langle x_1 \rangle_{R[\mathfrak{S}_\infty]}$. The results of [4] generalize this simple example and imply a surprising Noetherianity of the module $R$.

**Theorem 5.7.** *Let $I$ be a symmetric ideal of $R$. Then, $I$ is finitely generated as a module over $R[\mathfrak{S}_\infty]$. Moreover, there is finite set of polynomials $G$ such that for every polynomial $f$, we have $f \in I$ if and only if the polynomial reduction of $f$ with $G$ is $0$.*

The polynomial reduction appearing in Theorem 5.7 is a symmetric modification of the reduction in the context of normal (finite dimensional) polynomial rings.

**Example 5.8.** *The ideal $I = \langle x_1^3 x_3 + x_1^2 x_2^3, x_2^2 x_3^2 - x_2^2 x_1 + x_1 x_3^2 \rangle_{R[\mathfrak{S}_\infty]}$ from Problem 5.6 has a symmetric Gröbner basis given by:*

$$G = \mathfrak{S}_3 \cdot \{x_3 x_2 x_1^2, x_3^2 x_1 + x_2^4 x_1 - x_2^2 x_1, x_3 x_1^3, x_2 x_1^4, x_2^2 x_1^2\}.$$

*Once $G$ is found, testing whether a polynomial $f$ is in $I$ is computationally fast; for instance, one finds that $f \in I$ for the polynomial encountered in Problem 5.6.* □

My work with Aschenbrenner has focused on developing a theoretical framework for an algorithm we discovered that finds the set $G$ in Theorem 5.7. This involves a new and important partial order on monomials that respects the action of the symmetric group. We aim to make our techniques computationally effective, and we will apply them to the important finite dimensional situation. Many researchers in this field have been interested in incorporating our methods for computing Gröbner bases with symmetry because traditional techniques remove such structure. We also aim to generalize our results to other group actions and rings.

## 6. Gröbner Basis and Combinatorics

6.1. **Graphs and Commutative Algebra.** In recent years, it has been fruitful to study questions on graphs using commutative algebra. Let $G$ be a simple, undirected graph with vertex set $V = \{1, \ldots, n\}$ and edge set $E$. Fix a positive integer $k < n$, and let $C_k = \{c_1, \ldots, c_k\}$ be a $k$-element set. Each element of $C_k$ is called a *color*. A (vertex) *$k$-coloring* of $G$ is a map $\nu : V \to C_k$. We say that a $k$-coloring $\nu$ is *proper* if adjacent vertices receive different colors; otherwise $\nu$ is called *improper*. The graph $G$ is said to be *$k$-colorable* if there exists a proper $k$-coloring of $G$. Let $R = \mathbb{C}[x_1, \ldots, x_n]$, and consider the following ideals of $R$:

$$
\begin{aligned}
I_{n,k} &= \langle x_i^k - 1 : i \in V \rangle, \\
I_{G,k} &= I_{n,k} + \langle x_i^{k-1} + x_i^{k-2}x_j + \cdots + x_i x_j^{k-2} + x_j^{k-1} : \{i,j\} \in E \rangle.
\end{aligned}
$$

The zeroes of $I_{n,k}$ and $I_{G,k}$ represent $k$-colorings and proper $k$-colorings of the graph $G$, respectively. The idea of using roots of unity and ideal theory to study graph coloring problems seems to originate in Bayer's thesis [5], although it has appeared in many other places, including the work of de Loera [20] and Lovász [50]. These ideals are important because they allow for an algebraic formulation of $k$-colorability. Versions of the following theorem appeared in [2, 5, 20, 50, 53].

**Theorem 6.1.** *The following statements are equivalent:*

1. *The graph $G$ is not $k$-colorable.*
2. $\dim_{\mathbb{C}} R/I_{G,k} = 0$.
3. *The constant polynomial* $1$ *belongs to the ideal $I_{G,k}$.*
4. *The graph polynomial $f_G$ belongs to the ideal $I_{n,k}$.*

This theorem gives rise to algorithms [39] for determining $k$-colorability of a graph that are different from the traditional ones that use deletion and contraction. In [39], Windfeldt and I refined Theorem 6.1 and gave an algebraic characterization of uniquely colorable graphs. This provided us with algorithms to verify a counterexample of Akbari, Mirrokni, and Sadjad [1] to Xu's conjecture [71].

Independently, de Loera et al [21] have been studying complexity questions related to Gröbner bases and combinatorial optimization problems, such as graph colorings. The condition that $1 \in I_{G,k}$ can be checked by a Gröbner basis calculation, but the speed of this calculation is intimately related to the sizes of the degrees in a Nullstellensatz certificate. General theory says that this complexity is doubly exponential in the number of vertices $n$. However, for the special situation encountered here, there is much evidence to suggest that the complexity is only singly exponential. This would explain our experimental findings in [39]. Moreover, a careful implementation would allow for the computation of the chromatic numbers of large graphs, a significant advance.

I have begun working on this complexity problem with the team of de Loera, Margulies, and Woo. Our first approach will be to step through the papers of Sombra [70] and Kollár [44] for our special class of ideals. The hope is that some of the estimates in these works can be improved when the ideals come from graphs. We will also perform a similar inspection of the recent algorithmic advances on Castelnuovo-Mumford regularity [31], which is an important invariant measuring the complexity of an ideal. This theoretical work will be accompanied by a series of large-scale computations, which we will use to test our conjectures. In the process, we will develop a suite of tools that will be made available for other researchers working on symbolic computation and graph theory.

6.2. **Gröbner Bases and Partial Sums of Catalan Numbers.** The Casas-Alvero conjecture says that the following are equivalent for a degree $d$ monic polynomial $f \in \mathbb{C}[x]$:

(1) $f(x) = (x-b)^d$ for some $b \in \mathbb{C}$.
(2) $\gcd(f, \frac{d^i f}{dx^i}) \neq 1$ for all $k = 1, \ldots, d-1$.

For certain classes of degrees (for instance, prime powers), this result is known to be true [10]. Garcia and I have been studying this conjecture from the perspective of commutative algebra and Gröbner bases. Clearly (1) $\Rightarrow$ (2), and so the conjecture is (2) $\Rightarrow$ (1). Fix $d$ and let $r_1, \ldots, r_d$ be indeterminates. Also, set $f_d(x) = (x-r_1)\cdots(x-r_d)$. Consider the polynomials in $\mathbb{Z}[r_1, \ldots, r_d]$,

$$s_k = \mathrm{Res}(f_d, f_d^{(k)}), \quad k = 1, \ldots, d-1,$$

in which $f_d^{(k)}$ is the $k$th derivative of $f_d$ with respect to $x$. The conjecture may be reformulated in terms of the ideal $I_d = \langle s_k : k = 1, \ldots, d-1 \rangle$ and its variety:

**Conjecture 6.2.** *For the ideals $I_d$, we have $V(I_d) = \{(r, r, \ldots, r) \in \mathbb{C}^d : r \in \mathbb{C}\}$.*

Unfortunately, this ideal is very complicated [10], and so Garcia and I made a relaxation. Let $J_d$ be the ideals generated by the following polynomials:

$$t_k = \mathrm{Res}(x - r_1, f_d^{(k)}), \quad k = 1, \ldots, d-1.$$

In this case, it is readily verified that an analog of Conjecture 6.2 holds. We have some ideas for using the information gained from studying the ideals $J_d$. For instance, we hope to induct on the integer $l$ in a relaxation that replaces $(x - r_1)$ with $(x-r_1)\cdots(x-r_l)$ in the definition of $t_k$. The ideals $J_d$ then form the base case in this approach.

It turns out that the collection of $J_d$ are very interesting combinatorially. For instance, choosing the lexicographic ordering on monomials in $r_1, \ldots, r_d$ and computing the reduced Gröbner basis $G_d$ for each $J_d$, one finds that it consists of homogenous polynomials and that $|G_d|$ is identical (up to $d = 12$) to an interesting combinatorial sequence of numbers $\{1, 2, 4, 9, 23, 65, 197, 626, \ldots\}$, the partial sums of the Catalan numbers. Let $C_d = \frac{1}{d+1}\binom{2d}{d}$ denote the $d$th Catalan number. Then, we conjecture the following.

**Conjecture 6.3.** *For the Gröbner bases $G_d$, we have $|G_d| = \sum_{i=0}^{d-1} C_d$.*

We have much evidence for this conjecture. For instance, we have determined an algorithm that (conjecturally) generates the leading monomials in $G_d$. We also found a way to index the monomials generated by this algorithm, and we have proved that they are in combinatorial bijection with partial sums of Catalan numbers. The next step in our approach is to match the steps in this algorithm with the sequence of $S$-polynomial reductions that occur in a Gröbner basis calculation of the ideals $J_d$.

The phenomenon found in Conjecture 6.3 appears to be new, although Aval-Bergeron-Bergeron have also recently discovered an ideal having similar combinatorial structure that occurs naturally when computing the dimension of a quotient ring of quasisymmetric functions [6]. As in our case, they sought a bijection between the combinatorial objects they were studying and steps in a Gröbner basis calculation. Garcia and I also plan to investigate if there is a quotient ring of dimension $|G_d|$ hiding in our work.

## References


[1] S. Akbari, V. S. Mirrokni, B. S. Sadjad, *$K_r$-Free uniquely vertex colorable graphs with minimum possible edges.* Journal of Combinatorial Theory, Series B **82** (2001), 316–318.

[2] N. Alon, M. Tarsi, *Colorings and orientations of graphs.* Combinatorica **12** (1992), 125–134.

[3] M. Aschenbrenner and C. Hillar, *An algorithm for computing symmetric Gröbner bases in infinite dimensional polynomial rings*, in preparation.

[4] M. Aschenbrenner and C. Hillar, *Finite generation of symmetric ideals*, Trans. Amer. Math. Soc., **359** (2007), 5171–5192.

[5] D. Bayer, *The division algorithm and the Hilbert scheme.* Ph.D. Thesis, Harvard University, 1982.

[6] J.C. Aval, F. Bergeron, N. Bergeron, *Ideals of quasi-symmetric functions and super-covariant polynomials for $S_n$*, Adv. Math., **181** (2004), 353–367.

[7] D. Bessis, P. Moussa, M. Villani, *Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics*, J. Math. Phys. **16** (1975), 2318–2325.

[8] L. Blum, M. Shub, and S. Smale, "On a theory of computation and complexity over the real numbers," *Bull. Amer. Math. Soc.*, **21** (1989), no. 1, pp. 1–46.

[9] E. Candes, *Compressive sampling*, In Proceedings of the International Congress of Mathematicians (Madrid, 2006).

[10] H. von Bothmer, O. Labs, J. Schicho, C. van de Woestijne, *The Casas-Alvero conjecture for infinitely many degrees*, preprint.

[11] B. Buchberger, *An algorithmic criterion for the solvability of algebraic systems of equations*, Aequat. Math. **4** (1970), 374–383.

[12] S. Burgdorf, *Notes on $S_{m,4}(X^{\cdot}Y^2)$*, preprint.

[13] A. R. Camina, D. M. Evans, *Some infinite permutation modules*, Quart. J. Math. Oxford Ser. (2) **42** (1991), no. 165, 15–26.

[14] Ryan Canolty and Kai Miller. Large scale brain dynamics. NIPS Workshop, 2007.

[15] M. D. Choi, Z. D. Dai, T. Y. Lam, B. Reznick, *The Pythagoras number of some affine algebras and local algebras.* J. Reine Angew. Math. **336** (1982), 45–82.

[16] S. Cook, "The complexity of theorem proving procedures," *Proc. Annual ACM Symp. Theory Comput.* (STOC), **3** (1971), pp. 151–158.

[17] K. J. Le Couteur, *Representation of the function* $\mathrm{Tr}(\exp(A - \lambda B))$ *as a Laplace transform with positive weight and some matrix inequalities*, J. Phys. A: Math. Gen. **13** (1980), 3147–3159.

[18] K. J. Le Couteur, *Some problems of statistical mechanics and exponential operators*, pp 209–235 in Proceedings of the International Conference and Winter School of Frontiers of Theoretical Physics, eds. F. C. Auluck, L.S. Kothari, V.S. Nanda, Indian National Academy, New Dehli, 1977, Published by the Mac Millan Company of India, 1978.

[19] J. Daugman. Statistical Richness of Visual Phase Information: Update on Recognizing Persons by Iris Patterns. *International Journal of Computer Vision*, 45(1):25–38, 2001.

[20] J. A. de Loera, *Gröbner bases and graph colorings*, Beitrage zur Algebra und Geometrie **36** (1995), 89–96.

[21] J. A. de Loera, J. Lee, S. Margulies, S. Onn, *Expressing combinatorial optimization problems by systems of polynomial equations and the Nullstellensatz*, preprint.

[22] M. Drton, B. Sturmfels, S. Sullivant, *Algebraic factor analysis: Tetrads, pentads and beyond*, Probability Theory and Related Fields, to appear.

[23] M. Drmota, W. Schachermayer and J. Teichmann, *A hyper-geometric approach to the BMV-conjecture*, Monatshefte fur Mathematik **146** (2005), 179–201.

[24] M. Fannes and D. Petz, *Perturbation of Wigner matrices and a conjecture*, Proc. Amer. Math. Soc. **131** (2003), 1981–1988.

[25] M. Felsberg and G. Sommer. The monogenic signal. *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, 49(12):3136–3144, 2001.

[26] W.J. Freeman. *Mass Action in the Nervous System.* Academic Press, New York, 1975.

[27] P. Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10):474–480, 2005.

[28] D. Gondard, P. Ribenboim, *Le 17e problme de Hilbert pour les matrices*, Bull. Sci. Math., **98** (1974) 49–56.

[29] D. Hägele, *Proof of the cases $p \leq 7$ of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture*, J. Stat. Phys., to appear.

[30] F. Hansen, *Trace functions as Laplace transforms*, J. Math. Phys., **47** 043504 (2006).

[31] A. Hashemi, *Computation of Castelnuovo-Mumford regularity and satiety*, ISSAC 2007, preprint.

[32] J. Håstad, "Tensor rank is NP-complete," *J. Algorithms*, **11** (1990), no. 4, pp. 644–654.

[33] C. Hillar, *Sums of polynomial squares over totally real fields are rational sums of squares*, Proc. Amer. Math. Soc., **137** (2009), 921–930.

[34] C. Hillar, *Advances on the Bessis-Moussa-Villani Trace Conjecture*, Lin. Alg. Appl., **426** (2007), 130–142.

[35] C. Hillar and C. R. Johnson, *On the positivity of the coefficients of a certain polynomial defined by two positive definite matrices*, J. Stat. Phys., **118** (2005), 781–789.

[36] C. Hillar and C. R. Johnson, *Symmetric word equations in two positive definite letters*, Proc. Amer. Math. Soc., **132** (2004), 945–953.

[37] C. Hillar and C R. Johnson, *Positive eigenvalues of generalized words in two Hermitian positive definite matrices*, in: Novel Approaches to Hard Discrete Optimization (P. Pardalos and H. Wolkowicz, eds.), Fields Institute Communications **37** (2003), 111–122.

[38] C. Hillar, J. Nie, *An elementary and constructive solution to Hilbert's 17th problem for matrices*, Proc. Amer. Math. Soc., **136** (2008), 73–76.

[39] C. Hillar and T. Windfeldt, "Algebraic characterization of uniquely vertex colorable graphs," *J. Combin. Theory Ser. B*, **98** (2008), no. 2, pp. 400–414.

[40] C. R. Johnson and C. Hillar, *Eigenvalues of words in two positive definite letters*, SIAM J. Matrix Anal. Appl., **23** (2002), 916–928.

[41] R.M. Karp, "Reducibility among combinatorial problems," pp. 85–103, in R.E. Miller and J.W. Thatcher (Eds), *Complexity of computer computations*, Plenum, New York, NY, 1972.

[42] I. Klep and M. Schweighofer, *Sums of Hermitian squares and the BMV conjecture*, preprint.

[43] I. Klep and M. Schweighofer, *Connes' embedding conjecture and sums of hermitian squares*, preprint.

[44] J. Kollár, *Sharp effective Nullstellensatz.* J. Amer. Math. Soc. **1** (1988), 963–975.

[45] T. Y. Lam, *Introduction To Quadratic Forms Over Fields*, American Mathematical Society, 2004.

[46] E. Landau, *Uber die Darstellung definiter Funktionen durch Quadrate*, Math Ann., **62** (1906), 272–285.

[47] P. Landweber and E. Speer, *On D. Hägeles approach to the Bessis-Moussa-Villani conjecture*, preprint.

[48] E. H. Lieb and R. Seiringer, *Equivalent forms of the Bessis-Moussa-Villani conjecture*, J. Stat. Phys., **115** (2004), 185-190.

[49] L.A. Levin, "Universal sequential search problems," *Probl. Inf. Transm.*, **9** (1973) no. 3, pp. 265–266.

[50] L. Lovász, *Stable sets and polynomials.* Discrete Mathematics **124** (1994), 137–153.

[51] A. Mead, E. Ruch, A. Schönhofer, *Theory of chirality functions, generalized for molecules with chiral ligands.* Theor. Chim. Acta **29** (1973), 269–304.

[52] N. Miller, $3 \times 3$ *cases of the Bessis-Moussa-Villani conjecture*, Princeton University Senior Thesis, 2004.

[53] M. Mnuk, *On an algebraic description of colorability of planar graphs.* In Koji Nakagawa, editor, *Logic, Mathematics and Computer Science: Interactions. Proceedings of the Symposium in Honor of Bruno Buchberger's 60th Birthday.* RISC, Linz, Austria, October 20-22 (2002), 177–186.

[54] P. Moussa, *On the representation of* $\mathrm{Tr}\left(e^{A-\lambda B}\right)$ *as a Laplace transform*, Rev. Math. Phys. **12**, 621–655 (2000).

[55] J. Nie, K. Ranestad, B. Sturmfels, *The algebraic degree of semidefinite programming*, math.CO/0611562.

[56] B. Olshausen, D. Field, *Emergence of simple-cell receptive field properties by learning a sparse code for natural images*, Nature, 381: 607–609, 1996

[57] A. Papachristodoulou, P. A. Parrilo, S. Prajna, *Introducing SOSTOOLS: A General Purpose Sum of Squares Programming Solver*. Proceedings of the IEEE Conference on Decision and Control (CDC), Las Vegas, NV. 2002.

[58] A. Papachristodoulou, P. A. Parrilo, S. Prajna, *New Developments in Sum of Squares Optimization and SOSTOOLS*. Proceedings of the American Control Conference (ACC), Boston, MA. 2004.

[59] P. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*. Math. Program., Ser. B **96** (2003), 293–320.

[60] P. Parrilo, *Exploiting algebraic structure in sum of squares programs*, Positive polynomials in Control, Lecture Notes in Control and Information Sciences, Vol. 312, pp. 181–194, Springer, 2005.

[61] P. Parrilo, B. Sturmfels, *Minimizing polynomial functions*, Algorithmic and quantitative real algebraic geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 60, pp. 83–99, AMS.

[62] J. Portilla and E.P. Simoncelli. A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients. *International Journal of Computer Vision*, 40(1):49–70, 2000.

[63] Y. Pourchet, *Sur la représentation en somme de carrés des polynômes á une indéterminée sur un corps de nombres algébriques*, Acta Arith. **19** (1971), 89–104.

[64] A. Prestel, C. N. Delzell, *Positive Polynomials: From Hilbert's 17th Problem to Real Algebra*, Springer, 2001.

[65] Benjamin Rect, Maryam Fazel, and Pablo A. Parrilo, *Guaranteed minimum-rank solution of linear matrix equations via nuclear norm minimization.* (Preprint, 2007)

[66] E. Ruch, A. Schönhofer, *Theorie der Chiralitätsfunktionen*, Theor. Chim. Acta **19** (1970), 225–287.

[67] E. Ruch, A. Schönhofer, I. Ugi, *Die Vandermondesche Determinante als Näherungsansatz für eine Chiralitätsbeobachtung, ihre Verwendung in der Stereochemie und zur Berechnung der optischen Aktivität*, Theor. Chim. Acta **7** (1967), 420–432.

[68] M. Schweighofer, *Algorithmische Beweise für Nichtnegativ- und Positivstellensätze*, Diplomarbeit an der Universität Passau, 1999.

[69] T.J. Sejnowski and O. Paulsen. Network Oscillations: Emerging Computational Principles. *Journal of Neuroscience*, 26(6):1673, 2006.

[70] M. Sombra, *Bounds for the Hilbert function of polynomial ideals and for the degrees in the Nullstellensatz*. J. Pure Appl. Algebra **117** & **118** (1997) 565–599.

[71] S. Xu, *The size of uniquely colorable graphs*, Journal of Combinatorial Theory Series B **50** (1990), 319–320.