THE POLYNOMIAL EIGENVALUE PROBLEM IS WELL CONDITIONED FOR RANDOM INPUTS[*]

DIEGO ARMENTANO[†] AND CARLOS BELTRÁN[‡]

Abstract. We compute the exact expected value of the squared condition number for the polynomial eigenvalue problem, when the input matrices have entries coming from the standard complex Gaussian distribution, showing that in general this problem is quite well conditioned.

1. Introduction. Recall the homogeneous generalized eigenvalue problem (GEVP): given two $n \times n$ matrices A, B, find $(\alpha, \beta) \in \mathbb{C}^2 \setminus \{(0,0)\}$ such that

$$\det(\beta A - \alpha B) = 0.$$

The point (α, β) is called a generalized eigenvalue of (A, B) and the corresponding nonzero vectors $x, y \in \mathbb{C}^n$ satisfying

$$(\beta A - \alpha B)x = 0, \quad y^*(\beta A - \alpha B) = 0,$$

are called the right and left eigenvectors of (A, B).

The equations defining the generalized eigenvalues are homogeneous in (α, β) and x, y, respectively. This means that (α, β) is a generalized eigenvalue if and only if any nonzero multiple of (α, β) is also a generalized eigenvalue and similarly for x and y. It is thus natural to consider the eigenvalues and eigenvectors in projective spaces $\mathbb{P}(\mathbb{C}^2)$ and $\mathbb{P}(\mathbb{C}^n)$ where, for $b \geq 2$, $\mathbb{P}(\mathbb{C}^b)$ is the set of one-dimensional subspaces of $\mathbb{C}^b$ (see section 2).

The polynomial eigenvalue problem (PEVP), a well known generalization of the GEVP, is among the most important problems in numerical linear algebra; see the excellent survey [26] for context. The question now is: given $A_0, \ldots, A_d \in \mathbb{C}^{n \times n}$, find $(\alpha, \beta) \in \mathbb{P}(\mathbb{C}^2)$ such that

$$\mathcal{F}(A, (\alpha, \beta)) = \det(P(A, \alpha, \beta)) = 0,$$

where $A = (A_0, \ldots, A_d)$ and

$$(1.1) \quad P(A, \alpha, \beta) = \beta^d A_0 + \alpha\beta^{d-1} A_1 + \cdots + \alpha^{d-1}\beta A_{d-1} + \alpha^d A_d = \sum_{j=0}^{d} \alpha^j \beta^{d-j} A_j.$$

[†]Universidad de la Republica, Montevideo, 11400 (diego@cmat.edu.uy).

[‡]Departamento de Matemáticas, Estad. y Comput., Universidad de Cantabria, Santander, E-39071 (carlos.beltran@unican.es).

A solution $(\alpha, \beta) \in \mathbb{P}(\mathbb{C}^2)$ of $\mathcal{F}(A, (\alpha, \beta)) = 0$ is simply called an eigenvalue of A. Let us denote by $\mathrm{Eig}(A)$ the set of eigenvalues of A.

For generic input A, there exist nd such eigenvalues. Right and left eigenvectors of A are defined as in the GEVP case. If we set $d = 1$ we get (up to a sign convention) the GEVP.

A good deal of effort has been dedicated in the last years to find efficient algorithms for particular instances of the PEVP such as the quadratic eigenvalue problem; see [5, 25, 20] (the inverse version [24] is also of interest). The nowadays method of choice for solving the PEVP is to linearize it obtaining a GEVP (in dimension nd) that can be solved using standard eigenvalue solvers. Different ways to linearize the problem can be analyzed in search for optimal ones in terms of stability or numerical convenience; see [22].

A fundamental question regarding the numerical solution of the PEVP is the stability, governed by the so-called condition number [16, 21, 19]. When we consider the natural Hermitian structure over $\mathbb{P}(\mathbb{C}^2)$ and the space of matrices (see section 2), the condition number for the GEVP has a nice expression computed in [16, section 6]: for a matrix pair $(A, B) \in \mathbb{C}^{n \times n} \times \mathbb{C}^{n \times n}$ and an eigenvalue $(\alpha, \beta) \in \mathbb{P}(\mathbb{C}^2)$, we have

$$\mu((A, B), (\alpha, \beta)) = \frac{\|(\alpha, \beta)\| \, \|x\| \, \|y\|}{|\bar{\alpha} y^* A x + \bar{\beta} y^* B x|} \, \|(A, B)\|_F,$$

where x and y are the corresponding right and left eigenvectors, and $\| \cdot \|_F$ denotes Frobenius norm (the factor $\|(A, B)\|_F$ is missing in [16] since Dedieu and Tisseur compute the absolute condition number instead of the relative condition number). This formula is indeed a consequence of the definition of the condition number as the maximum of the change in the eigenvalue when the input is locally perturbed (see section 2 for a more formal definition).

The same definition as the maximum possible change in the eigenvalue when the input is perturbed is valid for the more general PEVP. An explicit formula for the condition number for the PEVP was derived in [16, Theorem 4.2],

$$(1.2) \qquad \mu(A, (\alpha, \beta)) = \left(\sum_{j=0}^{d} |\alpha|^{2j} |\beta|^{2d-2j} \right)^{1/2} \frac{\|x\| \, \|y\|}{|y^* v|} \, \|A\|_F,$$

where again $A = (A_0, \ldots, A_d)$, $(\alpha, \beta) \in \mathrm{Eig}(A)$, and x and y are the corresponding right and left eigenvectors and

$$(1.3) \qquad v = \bar{\beta} \frac{\partial}{\partial \alpha} P(A, \alpha, \beta) x - \bar{\alpha} \frac{\partial}{\partial \beta} P(A, \alpha, \beta) x.$$

Note that the condition number depends both on the input A, and on the particular eigenvalue (α, β). It is customary to consider

$$\mu_{max}(A) = \max_{(\alpha, \beta) \in \mathrm{Eig}(A)} \mu(A, (\alpha, \beta)).$$

Computing the condition number for any numerical problem (including the PEVP) is a time-consuming task that suffers from intrinsic stability problems as pointed out in [17]. It is hence usual to estimate average values of the condition number for a given family of inputs, in such a way that we can rely on probabilistic arguments instead of computing condition numbers of particular inputs. The linear algebra case is probably

the most studied one; see [18, 14] for estimates on Turing's condition number with different normalizations. Theoretical results describing the average condition number of the standard eigenvalue–eigenvector problem have been recently obtained [3, 2], but we are not aware of any previous similar result for the PEVP case. In this paper we fill this gap. Through this manuscript, we will denote by

$$\mathrm{E}_w(f(w)) \quad \text{or} \quad \mathrm{E}_{w \in W}(f(w))$$

the expected value of a function $f : W \to \mathbb{R}$, where $w \in W$ follows some prescribed distribution in the probability space W.

Recall that a complex random variable has the standard Gaussian distribution $\mathcal{N}_{\mathbb{C}}(0,1)$ when the real part and the imaginary part are independent and identically distributed real Gaussian centered random variables with variance $1/2$ (see section 2 for details). Then we have the following.

THEOREM 1.1. *Let* $A = (A_0, \ldots, A_d) \in \mathbb{C}^{(d+1)n^2}$ *be chosen at random with independent entries following the distribution* $\mathcal{N}_{\mathbb{C}}(0,1)$. *Then, the expected value of the squared condition number for the PEVP satisfies*

$$\mathrm{E}_A \left(\frac{1}{dn} \sum_{(\alpha,\beta) \in \mathrm{Eig}(A)} \mu(A, (\alpha, \beta))^2 \right) = \frac{(d+1)n^2 - 1}{d}.$$

A trivial consequence of Theorem 1.1 is the following.

COROLLARY 1.2. *Let* $A = (A_0, \ldots, A_d) \in \mathbb{C}^{(d+1)n^2}$ *be chosen as in Theorem* 1.1. *Then,*
$$\mathrm{E}_A \left(\mu_{max}(A)^2 \right) \leq ((d+1)n^2 - 1)n,$$

and in particular

$$\mathrm{E}_A \left(\log \left(\mu_{max}(A) \right) \right) \leq \frac{3}{2} \log n + \frac{1}{2} \log(d+1).$$

Since the number of digits needed to describe accurately the output of a numerical problem is controlled by the logarithm of the condition number, we can now see from Corollary 1.2 that there is no intrinsic obstruction for computing solutions to the PEVP, even for very large values of n and d.

The proof of Theorem 1.1 (see section 3.5) will be a consequence of a much more general result, Theorem 2.1. We also prove similar results for other numerical problems; see sections 3.1, 3.2, 3.3, 3.4, 3.6, 3.7 for other examples of application.

Remark 1.3. The real case (i.e., the case that the coefficients of the matrices involved are, say, obtained from the real standard Gaussian distribution) can also be analyzed. The techniques, however, are quite more involved and heavily use sophisticated results from real semialgebraic geometry. It turns out that these real random polynomial eigenvalue problems are also quite well conditioned on the average; see the main result in [7].

Remark 1.4. Since the condition number satisfies $\mu(tA, (\alpha, \beta)) = \mu(A, (\alpha, \beta))$ for any nonzero $t \in \mathbb{C}$, the results above also apply in the case that the vector of matrices $A = (A_0, \ldots, A_d)$ is distributed according to any centrally symmetric distribution in $\mathbb{C}^{(d+1)n^2}$. Note that this is different from the coefficients of the matrices following any central distribution!

Remark 1.5. One standard approach to solve a polynomial equation $p(x) = 0$ in one complex variable is to construct the companion matrix of the given polynomial and compute its eigenvalues (this is the method used by MATLAB's command `root`). It is usually convenient to balance the companion matrix before applying QR or some other standard eigensolver. Specific eigensolvers that make use of the special structure of companion matrices have also been developed; see [10, 9, 30, 4] and references therein. The stability of the process can be improved by using a companion matrix different from the standard one; see [15].

Currently existing eigenvalue solvers exhibit remarkable stability and accuracy properties, and indeed the above process shows excellent practical performance. In contrast, computing eigenvalues of a matrix by writing down its characteristic polynomial and solving it by Newton's method (or other basic procedure) shows quite a poor performance in practice. This may have contributed to support the idea, extended in some part of the community, that polynomial root finding is poorly conditioned, while eigenvalue computation is well conditioned. Indeed, in [29, page 92] root finding is presented as a classic example of ill-conditioned problem and so is the computation of eigenvalues in the non-symmetric case.

Theorem 1.1 (with $n = 1$ for the polynomial case) shows that this is not the case at least if the coefficients are drawn from $\mathcal{N}_{\mathbb{C}}(0,1)$:

> *Both polynomial root finding and polynomial eigenvalue problems are pretty well conditioned on the average, and indeed polynomial root finding has average squared condition number exactly equal to 1.*

An explanation of the poor practical performance of the characteristic polynomial method for computing eigenvalues is given in [13].

Remark 1.6. The condition number in a numerical problem measures the change in the solution when the input is locally perturbed, giving a worst-case bound: the worst (local) perturbation in the input defines the condition number of a problem.

In some situations, perturbations of the input can be expected to be "random" in the sense that there is no preferred direction. Of course, in this case the observed condition number of the problem may seem better than the theoretical bound, since most perturbations will be quasiorthogonal to the worst direction from the concentration of measure phenomenon. The condition number for random direction in the perturbation was called in [1] the *stochastic condition number* μ_{st}. Using [16, Theorem 4.2], its formula for the PEVP is given by

$$\mu_{\mathrm{st}}(A, (\alpha, \beta))^2 = \int_{\dot{A} \in \mathbb{S}} \frac{|y^* P(\dot{A}, \alpha, \beta) x|^2}{|y^* v|^2} \, \|A\|_F^2 \, d\sigma_{\mathbb{S}}(\dot{A}),$$

where $\mathbb{S}$ denotes the unit sphere of $\mathbb{C}^{n^2(d+1)}$, with the rotationally invariant probability measure $d\sigma_{\mathbb{S}}$, and v is given by (1.3). From [1, Theorem 1] we have $\mu^2 = (d+1)n^2\mu_{\mathrm{st}}^2$. Then, by Theorem 1.1 we obtain the equality

$$(1.4) \qquad \mathrm{E}_A \left(\frac{1}{dn} \sum_{(\alpha, \beta) \in \mathrm{Eig}(A)} \mu_{\mathrm{st}}(A, (\alpha, \beta))^2 \right) = \frac{(d+1)n^2 - 1}{d(d+1)n^2} \approx \frac{1}{d},$$

where $A \in \mathbb{C}^{(d+1)n^2}$ is chosen at random with independent $\mathcal{N}_{\mathbb{C}}(0,1)$ entries.

Thus, for a random eigenvalue of a random input and a random direction in the perturbation, one should expect the squared relative change in the solution to be around $\frac{(d+1)n^2-1}{d(d+1)n^2}$ times the size of the perturbation (relative to $\|A\|_F = \|(A_0, \ldots,$
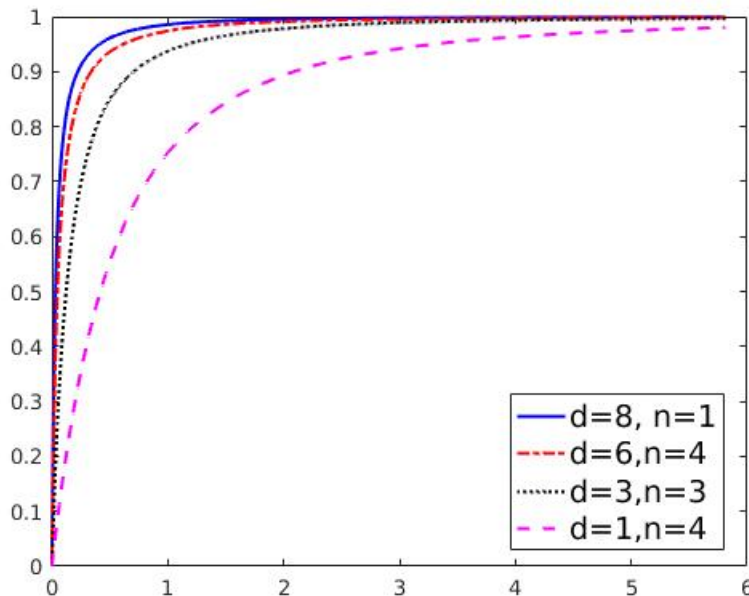
FIG. 1.1. *Empirical cdfs of the perturbation in a randomly chosen eigenvalue of a random problem instance for different values of the parameters d, n. Our theoretical expected values from (1.4) are* $0.1111, 0.1652, 0.3241, 0.9688$ *from top to bottom. The computed (empirical) expected values are* $0.1106, 0.1633, 0.3312, 0.9593$.

$A_d)\|_F$). In Figure 1.1 we have computed the empirical cumulative distribution function (cdf) of the size of that perturbation for different choices of d, n. For each of these choices we have computed $100,000$ random problem instances and corresponding random perturbation directions and then we have computed the distance from the eigenvalue of the unperturbed problem to that of the perturbed problem (the size of the perturbation is 10^{-6} relative to the norm of the input). The computations have been made with MATLAB's command polyeig.

2. Geometric framework: A general result. We will follow the path of [27, 16, 28] but considering a general setting of input and output spaces that can be applied to many different situations (see, for example, [11] or [1]).

2.1. The complex projective space and its subvarieties. For an integer $b \geq 2$, the projective space $\mathbb{P}(\mathbb{C}^b)$ is a classical object defined as the set of complex lines through 0 in $\mathbb{C}^b$, that is, as the set of possible directions in $\mathbb{C}^b$. A detailed yet accessible description of the geometry of $\mathbb{P}(\mathbb{C}^b)$ can be found in [11, page 190], for example. An algebraic subvariety Z of $\mathbb{P}(\mathbb{C}^b)$ is the subset of $\mathbb{P}(\mathbb{C}^b)$ given as the zero set of some collection of homogeneous polynomial equations, and it is irreducible if it cannot be decomposed into two proper subvarieties of Z or, equivalently, if the smooth points of Z—the points at which Z is a complex manifold—form a connected manifold. See, for example, [23] for a classical approach to (complex, projective) algebraic varieties.

The degree of an algebraic variety is defined in geometric terms: if an algebraic variety $Z \subseteq \mathbb{P}(\mathbb{C}^b)$ is of dimension m, then we can intersect it with a linear subspace L of $\mathbb{P}(\mathbb{C}^b)$ of codimension m and expect the number of intersection points $\#(Z \cap L)$

to be finite. It turns out that such number is finite and constant for a *generic* linear subspace, and it is called the degree of the variety; see [23, page 415]. This is a fundamental difference between real and complex algebraic geometry.

An important property of algebraic subvarieties of $\mathbb{P}(\mathbb{C}^b)$ is the rigidity of their volume: if an algebraic variety $Z \subseteq \mathbb{P}(\mathbb{C}^b)$ has dimension m and degree d then its volume (w.r.t. the natural metric in $\mathbb{P}(\mathbb{C}^b)$, the Fubini–Study metric defined in section 2.3 below) is exactly equal to d times the volume of the complex projective space of the same dimension as Z. In other words,

$$(2.1) \qquad Vol(Z) = vol(\mathbb{P}(\mathbb{C}^{m+1}))\, d = \frac{\pi^m}{\Gamma(m+1)} d;$$

see [12, Corollary 20.10].

2.2. The input-output spaces and solution variety. In many numerical problems there is a space of inputs $\mathcal{I}$ that we can identify with an m-dimensional complex vector space, and a space of outputs $\mathcal{O}$ that we can identify with a one-dimensional projective irreducible algebraic subvariety of $\mathbb{P}(\mathbb{C}^b)$, for some $b \in \mathbb{N}$. Some concrete examples are discussed below.

We denote by $d_{\mathcal{O}}$ the degree of $\mathcal{O}$; that is, the generic number of intersections of $\mathcal{O}$ with a linear subspace in $\mathbb{P}(\mathbb{C}^b)$ of codimension 1. We denote elements in $\mathcal{I}$ by p and elements in $\mathcal{O}$ by z. Under the identification $\mathcal{I} = \mathbb{C}^m$, we consider a polynomial $\mathcal{F} : \mathbb{C}^m \times \mathbb{C}^b \to \mathbb{C}$,

$$\mathcal{F}(p, z)$$

bihomogeneous in its two variables with degrees

$$\deg_p \mathcal{F} = r, \quad \deg_z \mathcal{F} = s.$$

We look at the problem:

given $p \in \mathcal{I}$, find $z \in \mathcal{O}$ such that $\mathcal{F}(p, z) = 0$.

Then, we consider the *solution variety*

$$\mathcal{V} = \{(p, z) \in \mathcal{I} \times \mathcal{O} : \mathcal{F}(p, z) = 0\} \subseteq \mathcal{I} \times \mathbb{P}(\mathbb{C}^b),$$

as well as the two natural projections $\pi_1 : \mathcal{V} \to \mathcal{I}$, $\pi_2 : \mathcal{V} \to \mathcal{O}$.

Here are some basic examples of the general setting above, including the PEVP case:

A $\mathcal{I} \equiv \mathbb{C}^{d+1}$ is the set of polynomials of degree at most d, homogeneous in two variables, and we search for a zero $z \in \mathbb{P}(\mathbb{C}^2)$. Here, $\mathcal{F}$ is the evaluation map

$$\mathcal{F}(p, z) = p(z).$$

B $\mathcal{I} \equiv \mathbb{C}^{n \times n} \times \mathbb{C}^{n \times n} \equiv \mathbb{C}^{2n^2}$ is the set of pairs of matrices (A, B), and we search for a generalized eigenvalue $(\alpha, \beta) \in \mathbb{P}(\mathbb{C}^2)$, i.e., a solution of

$$\mathcal{F}((A, B), (\alpha, \beta)) = \det(\beta A - \alpha B) = 0.$$

In general, there exist n generalized eigenvalues.

C $\mathcal{I} \equiv \mathbb{C}^{n \times n} \times \cdots \times \mathbb{C}^{n \times n} \equiv \mathbb{C}^{(d+1)n^2}$ (there are $d+1$ copies of $\mathbb{C}^{n \times n}$), and we want to solve the PEVP; that is, given $A = (A_0, \ldots, A_d)$ we want to find $(\alpha, \beta) \in \mathbb{P}(\mathbb{C}^2)$ such that

$$\mathcal{F}(A, (\alpha, \beta)) = \det(P(A, \alpha, \beta)) = 0,$$

where P is given in (1.1). In general, there exist dn eigenvalues. Note that A and B can be seen as a particular cases of C.

D One can also consider structured, sparse, or lacunary versions of the problems above. For example, the same problems where some of the entries of the matrices are set to 0 or where the matrices are Toeplitz or stochastic or have some other structure.

2.3. Hermitian products, norms, and Riemannian structures. The space $\mathbb{C}^b$ is equipped with the canonical Hermitian inner product $\langle \cdot, \cdot \rangle$. The associated projective space $\mathbb{P}(\mathbb{C}^b)$ is a smooth manifold which carries a natural Riemannian metric, namely, the real part of the Fubini–Study metric. This metric is the Hermitian structure on $\mathbb{P}(\mathbb{C}^b)$, given in the following way: for $z \in \mathbb{C}^b$,

$$(2.2) \qquad \langle w, w' \rangle_z := \frac{\langle w, w' \rangle}{\|z\|^2}$$

for all w, w' in the Hermitian complement $z^\perp$ of z in $\mathbb{C}^b$. (See [11, page 190] for example).

If we are to define condition numbers and to measure properties of random inputs, it is customary to endowe the input space $\mathcal{I}$ with a Hermitian inner product $\langle \cdot, \cdot \rangle_{\mathcal{I}}$. We denote by $\|\cdot\|_{\mathcal{I}}$ the induced norm, and we identify $\mathcal{I} \equiv \mathbb{C}^m$ as follows: choose any basis $p_1, \ldots, p_m$ orthonormal w.r.t $\langle \cdot, \cdot \rangle_{\mathcal{I}}$ and consider the linear isometry

$$(2.3) \qquad \begin{matrix} \mathbb{C}^m & \to & \mathcal{I} \\ \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} & \mapsto & \sum_{i=1}^{m} x_i p_i. \end{matrix}$$

Any such isometry (i.e., any choice of orthonormal basis of $\mathcal{I}$) gives a valid identification $\mathcal{I} \equiv \mathbb{C}^m$.

A random choice of $p \in \mathcal{I}$ is to be understood as a random variable for the Gaussian density on the complex vector space $\mathcal{I}$ given by

$$(2.4) \qquad \frac{1}{\pi^m} e^{-\|p\|_{\mathcal{I}}^2}, \quad p \in \mathcal{I}.$$

In other words, under the equivalence $\mathcal{I} \equiv \mathbb{C}^m$, a random element $p \in \mathcal{I}$ is obtained by generating a random element $x \in \mathbb{C}^m$ with the standard complex Gaussian distribution $\mathcal{N}_{\mathbb{C}}(0, \mathrm{Id}_m)$ on $\mathbb{C}^m$. Thus, generating $x_1, \ldots, x_m$ independently and identically distributed with density $\pi^{-1} e^{-|\zeta|^2}$, $\zeta \in \mathbb{C}$, a random element in $\mathcal{I}$ is given by $\sum_{i=1}^{m} x_i p_i$. A key feature of the Gaussian distribution is that it is independent of the particular choice of the basis $p_1, \ldots, p_m$.

We will use the following notation for norms.
- By $\|\cdot\|$ we mean the standard Hermitian vector norm. In particular, if A is a matrix, we have $\|A\| = \|A\|_F$ (Frobenius norm).
- By $\|A\|_2$ (with A a matrix or a linear operator) we denote the operator 2-norm of A.
- By $\|\cdot\|_z$ we mean the norm in the tangent space $T_z \mathbb{P}(\mathbb{C}^b)$; that is, the norm induced by the Hermitian product in (2.2).

2.4. The geometric definition of the condition number. The relative condition number of numerical problems in this generic input-output setting is defined as follows (see [11], [1], or [12, section 14.1] for a general setting, or [16] for a specification to the PEVP). Let z_0 be a solution for an input p_0. Assume that z_0 is a smooth point of $\mathcal{O}$, and let $\dot{z}_0$ be any nonzero vector in $T_{z_0}\mathcal{O}$. When the directional derivative

$$DF(p_0, z_0)(0, \dot{z}_0)$$

is not equal to 0, from the implicit function theorem, there is a locally defined *solution map*, denoted by S*ol*, which sends an input p (close to p_0) to its output z (close to z_0). This map is given by the composition $\pi_2 \circ \pi_1^{-1}$, locally defined in a neighborhood of p_0. The condition number at (p_0, z_0) is then defined by the operator norm of the derivative of the solution map, normalized by the norm of the input

$$(2.5) \qquad \mu(\mathcal{F}, p_0, z_0) = \|p_0\|_{\mathcal{I}} \|DSol(p_0, z_0)\|_2 = \sup_{\dot{p} \in \mathcal{I}} \frac{\|\dot{z}\|/\|z_0\|}{\|\dot{p}\|_{\mathcal{I}}/\|p_0\|_{\mathcal{I}}},$$

where $\dot{z} = DSol(p_0, z_0)\dot{p}$. (We will write $\mu(p_0, z_0)$ when the mapping $\mathcal{F}$ is clear from the context.)

In the case that S*ol* is just Lipschitz, the condition number uses the local Lipschitz constant instead of the 2-norm of the derivative, but we will not deal with this case here. The condition number is set to ∞ if $DF(p_0, z_0)(0, \dot{z}_0) = 0$ for some (i.e., for all) nonzero vector $\dot{z}_0 \in T_{z_0}\mathcal{O}$, or if z_0 is not a smooth point of $\mathcal{O}$.

This geometric definition of the condition number is inspired by its intuitive definition: it is a local bound for the change in the output under perturbations on the input, both measured in relative error terms. The classical condition number of linear algebra $\kappa(A) = \|A\|_2 \|A^{-1}\|_2$ does not exactly follow this definition since in our general framework we measure the relative error with respect to the norm of the input as a vector. Indeed, our definition gives the so called Demmel's condition number,

$$\mu(A, x) = \|A\|_F \|A^{-1}\|_2,$$

for the problem of solving $Ax = b$ with b fixed.

2.5. Main result: General version. Our main result is a general formula for the expected value of the condition number. Its main feature is that it is valid for the general input-output setting described above. For $p \in \mathcal{I}$ and $z \in \mathcal{O}$, let $n(p) = \#\pi_2(\pi_1^{-1}(p))$ be the number of z's such that $\mathcal{F}(p, z) = 0$, and $\mathcal{V}_z = \pi_1(\pi_2^{-1}(z))$, i.e.,

$$n(p) = \#\{z \in \mathcal{O} : \mathcal{F}(p, z) = 0\}, \quad \mathcal{V}_z = \{p \in \mathcal{I} : \mathcal{F}(p, z) = 0\},$$

where we set $n(p) = +\infty$ in case that $\pi_1^{-1}(p)$ is an infinite set.

THEOREM 2.1. *With the notations above, assume that:*
- *$n(p)$ is finite (equivalently, $\pi_2(\pi_1^{-1}(p)) \neq \mathcal{O}$ since these are algebraic sets and $\mathcal{O}$ is irreducible) for all $p \in \mathcal{I}$ in the complement of some $m - 2$ dimensional subvariety, and there exists $p_0 \in \mathcal{I}$ such that*

$$(2.6) \qquad\qquad n(p_0) = sd_{\mathcal{O}},$$

  where, recall, $s = \deg_z \mathcal{F} \in \mathbb{N}$.
- *$\mathcal{V}_z$ is an $m - 1$ dimensional variety for all $z \in \mathcal{O}$, and there exists $z_0 \in \mathcal{O}$ such that the degree of $\mathcal{V}_{z_0}$ is equal to $r = \deg_p \mathcal{F}$.*

Then, for all $p \in \mathcal{I}$ in the complement of some zero-measure set, the equation $\mathcal{F}(p, z) = 0$ has exactly $sd_{\mathcal{O}}$ solutions in $\mathcal{O}$. Moreover, the expected squared condition number, for the standard Gaussian distribution on $\mathcal{I}$ given by (2.4), satisfies

$$(2.7) \qquad \mathrm{E}_{p \in \mathcal{I}} \left(\frac{1}{sd_{\mathcal{O}}} \sum_{z : \mathcal{F}(p,z)=0} \mu(p,z)^2 \right) = \frac{(m-1)r}{s}.$$

Remark 2.2. Since the hypotheses of Theorem 2.1 only involve the algebraic structure of the problem, we conclude that the expected square condition number for the standard Gaussian does not depend on the chosen Hermitian inner product on $\mathcal{I}$ (see Examples 3.1 and 3.2).

2.6. Proof of Theorem 2.1. It will be helpful to recall the notion of a *constructible subset* of an algebraic variety Z (see [23, page 393]): it is a set of the form

$$\bigcup_{i=1}^{r} (V_i \setminus W_i),$$

where V_i, W_i are algebraic subvarieties of Z. In other words, a constructible set is a finite union of quasialgebraic varieties. Chevalley's theorem [23, page 395] asserts that the projection of a constructible set is a constructible set.

We organize the proof of Theorem 2.1 in several claims. The key step is Claim 5 which uses the smooth coarea formula (see Appendix B below). This result requires that the projections π_1, π_2 from the solution variety $\mathcal{V}$ are surjective submersions, and Claims 1 to 4 are just devoted to prove that these hypotheses hold once we remove lower dimensional algebraic subvarieties from $\mathcal{I}, \mathcal{O}$, and $\mathcal{V}$.

Claim 1: *There exists a proper subvariety $\mathcal{S} \subseteq \mathcal{I}$ such that for $p \in \mathcal{I} \setminus \mathcal{S}$ we have $n(p) = sd_{\mathcal{O}}$, and all the solutions $z \in \mathcal{O}$ of $\mathcal{F}(p, z) = 0$ are smooth points of $\mathcal{O}$ and have multiplicity equal to 1. In particular, $\mathcal{S}$ has Lebesgue measure zero on $\mathcal{I}$. Similarly, there exists at most a finite collection of $z \in \mathcal{O}$ such that the degree of $\mathcal{V}_z$ is different from r.* This is a classical fact, but we include a short proof using tools from classical complex analytic geometry since the same tools will be used later. The set

$$\hat{\mathcal{I}} = \{ p \in \mathcal{I} : n(p) = sd_{\mathcal{O}} \}$$

is constructible (see [23, page 398]). From (2.6), $\hat{\mathcal{I}} \neq \emptyset$, and furthermore, if $p \in \hat{\mathcal{I}}$, from Bézout's theorem [23, page 432], all the solutions $z \in \mathcal{O}$ of $\mathcal{F}(p, z) = 0$ have multiplicity equal to 1 (in particular, all z are smooth points of $\mathcal{O}$). Then, from the implicit function theorem for all $\widetilde{p}$ sufficiently close to p we also have $\widetilde{p} \in \hat{\mathcal{I}}$. In other words, $\hat{\mathcal{I}}$ is a constructible open set, thus a quasialgebraic set from [23, page 394]. In particular its complement,

$$\mathcal{S} = \mathcal{I} \setminus \hat{\mathcal{I}},$$

is an algebraic subvariety of $\mathcal{I}$, and thus its Lebesgue measure is zero. The first part of the claim follows. The second claim is proved in a similar way: the set of $z \in \mathcal{O}$ such that $\mathcal{V}_z$ has degree equal to r is open and constructible, thus quasialgebraic in $\mathcal{O}$, and hence its complement is an algebraic subset (a finite collection of points) of $\mathcal{O}$.

Claim 2: *For all but a finite set of $z \in \mathcal{O}$, the set*

$$\mathcal{S}_z = \{p \in \mathcal{S} : \mathcal{F}(p,z) = 0\}$$

is an algebraic variety of dimension at most $m - 2$. Assume that there is an infinite collection of points z such that $\mathcal{S}_z$ has dimension $m - 1$. Since the number of irreducible components of the $m - 1$ dimensional variety $\mathcal{S}$ is finite (recall that any algebraic variety can be written as a finite union of irreducible varieties; see, for example, [23, page 355]), this implies that there exists one such irreducible component Y of dimension $m - 1$ such that $\mathcal{S}_z \supseteq Y$ for an infinite collection of $z \in \mathcal{O}$. But then for all $p \in Y$ the number of solutions of $\mathcal{F}(p,z) = 0$ is infinite, that is, a contradiction to the hypotheses of the theorem.

Claim 3: *Let $\hat{\mathcal{O}}$ be the smooth part of $\mathcal{O}$, and let $\hat{\mathcal{V}} = \{(p,z) \in \hat{\mathcal{I}} \times \hat{O} : \mathcal{F}(p,z) = 0\}$. Then, π_1 restricted to $\hat{\mathcal{V}}$ is a submersion onto $\hat{\mathcal{I}}$.* Note that $\hat{O}$ is a complex 1-dimensional smooth submanifold of $\mathcal{O}$. From Claim 1, given $(p,z) \in \hat{\mathcal{V}}$, z is a solution to $\mathcal{F}(p,z) = 0$ of multiplicity 1, that is, $\frac{\partial}{\partial z}\mathcal{F}(p,z) \neq 0$. From the implicit function theorem $\hat{\mathcal{V}}$ is an m-dimensional complex manifold, its tangent space is

$$T_{(p,z)}\hat{\mathcal{V}} = \mathrm{Ker} D\mathcal{F}(p,z) \subseteq \mathcal{I} \times T_z \mathbb{P}(\mathbb{C}^b) \equiv \mathcal{I} \times z^\perp$$

and the projection $\pi_1 \mid_{\hat{\mathcal{V}}} : \hat{\mathcal{V}} \to \hat{\mathcal{I}}$ is a submersion at (p,z).

Claim 4: *After removing from $\hat{\mathcal{V}}$ at most a $(m-1)$-dimensional subvariety, the restriction of π_2 to $\hat{\mathcal{V}}$ is a submersion onto the complement of a finite set of points in $\hat{\mathcal{O}}$.* Note that $\pi_2 : \mathcal{V} \to \mathcal{O}$ is surjective by the hypotheses of the theorem. Now, from Claim 2 we have that for all but a finite number of points $z \in \hat{\mathcal{O}}$ there is some $p \in \mathcal{I}$ such that $p \in \mathcal{V}_z \setminus \mathcal{S}_z$, and hence the image of $\phi_2 := \pi_2 \mid_{\hat{\mathcal{V}}}$ contains $\hat{\mathcal{O}}$ except at most a finite number of points. Now, from Sard's theorem the set of singular values of ϕ_2 has zero-measure w.r.t. the Riemannian measure inherited from the Fubini–Study metric on $\mathbb{P}(\mathbb{C}^b)$. Furthermore, it is the projection of the constructible set consisting of the points $(p,z) \in \hat{V}$ such that $\frac{\partial}{\partial p}\mathcal{F}(p,z) = 0$. Hence, it is a zero-measure constructible set, that is, a finite collection of points in $\mathcal{O}$. The $(m-1)$-dimensional subvariety to be removed from $\hat{\mathcal{V}}$ is precisely the preimage of that finite set, which is by the hypotheses of the theorem a finite union of $(m-1)$-dimensional varieties, that is, itself a $(m-1)$-dimensional variety.

Claim 5: *For any measurable mapping $g : \mathcal{V} \to [0, \infty)$,*

$$(2.8) \quad \int_{p \in \mathcal{I}} \sum_{z \in \mathcal{O} : \mathcal{F}(p,z)=0} g(p,z)\,dp = \int_{z \in \mathcal{O}} \int_{p \in \mathcal{V}_z} \frac{NJ\pi_1(p,z)}{NJ\pi_2(p,z)} g(p,z)\,dp\,dz,$$

where NJ means Normal Jacobian (see Appendix B). In particular, these integrals are well defined. This is a consequence of the smooth coarea formula (Theorem B.1 below) applied to the two projections $\pi_1 : \hat{\mathcal{V}} \to \hat{\mathcal{I}}$ and $\pi_2 : \hat{\mathcal{V}} \to \hat{\mathcal{O}}$. From Claim 3, $\pi_1 \mid_{\hat{\mathcal{V}}}$ is a surjective submersion and we have

$$\int_{(p,z) \in \hat{\mathcal{V}}} NJ\pi_1(p,z)\,g(p,z)\,d(p,z) = \int_{p \in \hat{\mathcal{I}}} \sum_{z \in \mathcal{O} : \mathcal{F}(p,z)=0} g(p,z)\,dp.$$

On the other hand, from Claim 4, $\pi_2 \mid_{\hat{\mathcal{V}}}$ is a surjective submersion after removing some zero measure sets out of $\mathcal{V}$ and $\mathcal{O}$. We then have

$$\int_{(p,z)\in\hat{\mathcal{V}}} NJ\pi_1(p,z)\, g(p,z)\, d(p,z) = \int_{z\in\hat{\mathcal{O}}} \int_{p\in\hat{\mathcal{I}}:\mathcal{F}(p,z)=0} \frac{NJ\pi_1(p,z)}{NJ\pi_2(p,z)} g(p,z)\, dp.$$

Putting these two equalities together we have

$$\int_{p\in\hat{\mathcal{I}}} \sum_{z\in\mathcal{O}:\mathcal{F}(p,z)=0} g(p,z)\, dp = \int_{z\in\hat{\mathcal{O}}} \int_{p\in\hat{\mathcal{I}}:\mathcal{F}(p,z)=0} \frac{NJ\pi_1(p,z)}{NJ\pi_2(p,z)} g(p,z)\, dp.$$

Finally, from Claims 1 and 2 we can substitute $\hat{\mathcal{I}}$ and $\hat{\mathcal{O}}$ by $\mathcal{I}$ and $\mathcal{O}$, respectively, in this last equality, and the claim follows. Note that in order to change $\hat{\mathcal{I}}$ to $\mathcal{I}$ in the domain of the inner integral we need to use Claim 2.

Claim 6: *The quotient of the Normal Jacobians satisfies*

$$\frac{NJ\pi_1(p,z)}{NJ\pi_2(p,z)} \mu(p,z)^2 = \|p\|_{\mathcal{I}}^2, \quad (p,z) \in \hat{\mathcal{V}}.$$

Given $(p,z) \in \hat{\mathcal{V}}$, let $D\pi_2(p,z)^* : T_z\hat{\mathcal{O}} \to T_{(p,z)}\hat{\mathcal{V}}$ be the adjoint operator to $D\pi_2(p,z)$, and let

$$(\dot{p}_0, \dot{z}_0) = D\pi_2(p,z)^*(\dot{z}_0) \in T_{(p,z)}\hat{\mathcal{V}},$$

where $\dot{z}_0$ is any nonzero vector in $T_z\hat{\mathcal{O}}$, and let $(\dot{p}_j, \dot{z}_j), 1 \le j \le m-1$ be an orthonormal basis of the orthogonal complement of $(\dot{p}_0, \dot{z}_0)$ in $T_{(p,z)}\hat{\mathcal{V}}$ (this tangent space was described in Claim 3). Since the image of $D\pi_2(p,z)^*$ is the orthogonal complement to the kernel of $D\pi_2(p,z)$, then we have $\dot{z}_j = 0$ and $\dot{p}_j$ orthogonal to $\dot{p}_0$ for $1 \le j \le m-1$. Then, writing the linear operators $D\pi_1(p,z)$ and $D\pi_2(p,z)$ on this basis we get

$$NJ\pi_1(p,z) = |\det(D\pi_1(p,z))|^2 = \frac{\|\dot{p}_0\|_{\mathcal{I}}^2}{\|\dot{p}_0\|_{\mathcal{I}}^2 + \|\dot{z}_0\|_z^2}.$$

Similarly, we have

$$NJ\pi_2(p,z) = \frac{\|\dot{z}_0\|_z^2}{\|\dot{p}_0\|_{\mathcal{I}}^2 + \|\dot{z}_0\|_z^2}.$$

In particular,

$$\frac{NJ\pi_1(p,z)}{NJ\pi_2(p,z)} = \frac{\|\dot{p}_0\|_{\mathcal{I}}^2}{\|\dot{z}_0\|_z^2}.$$

Now, note from the definition of the condition number that $\mu(p,z)$ is precisely $\|p\|_{\mathcal{I}}$ times the operator norm of the linear operator $D\mathrm{Sol}(p,z)$ given by

$$D\pi_2(p,z) \circ D\pi_1^{-1}(p,z) : T_p\hat{\mathcal{I}} \to T_z\hat{\mathcal{O}}.$$

Since we have an orthogonal basis $\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_{m-1}$ satisfying

$$D\pi_2(p,z) \circ D\pi_1^{-1}(p,z)(\dot{p}_j) = 0, \quad 1 \le j \le m-1,$$

we conclude that

$$\frac{\mu(p,z)}{\|p\|_{\mathcal{I}}} = \|D\pi_2(p,z) \circ D\pi_1^{-1}(p,z)\|_2$$

$$= \frac{\|D\pi_2(p,z) \circ D\pi_1^{-1}(p,z)(\dot{p}_0)\|_z}{\|\dot{p}_0\|_{\mathcal{I}}} = \frac{\|D\pi_2(p,z)(\dot{p}_0, \dot{z}_0)\|_z}{\|\dot{p}_0\|_{\mathcal{I}}} = \frac{\|\dot{z}_0\|_z}{\|\dot{p}_0\|_{\mathcal{I}}}.$$

We have thus proved the equality in the claim.

Claim 7: *The following equality holds:*

$$\mathrm{E}_{p \in \mathcal{I}} \left(\sum_{z \in \mathcal{O}: \mathcal{F}(p,z)=0} \mu(p,z)^2 \right) = (m-1)\, r\, d_{\mathcal{O}}.$$

(And this readily implies the theorem). In order to prove Claim 7 we use Claims 5 and 6 with $g(p,z) = e^{-\|p\|_{\mathcal{I}}^2} \mu(p,z)^2 / \pi^m$, getting

$$\mathrm{E}_{p \in \mathcal{I}} \left(\sum_{z \in \mathcal{O}: \mathcal{F}(p,z)=0} \mu(p,z)^2 \right) = \frac{1}{\pi^m} \int_{z \in \mathcal{O}} \int_{p \in \mathcal{V}_z} \|p\|_{\mathcal{I}}^2 e^{-\|p\|_{\mathcal{I}}^2}\, dp\, dz.$$

Now, for any $z \in \mathcal{O}$ (except at most a finite number) using Lemma A.1 (recall, we have identified $\mathcal{I}$ with $\mathbb{C}^m$ through a linear isometry (2.3)) we have

$$(2.9) \qquad \int_{p \in \mathcal{V}_z} \|p\|_{\mathcal{I}}^2 e^{-\|p\|_{\mathcal{I}}^2}\, dp = \pi^{m-1}\, r\, (m-1),$$

since the set $\mathcal{V}_z$ is (for almost all $z \in \mathcal{O}$) an algebraic subvariety of dimension $m-1$ and degree r. Using that the volume of $\mathcal{O}$ is $d_{\mathcal{O}} Vol(\mathbb{P}(\mathbb{C}^2)) = \pi d_{\mathcal{O}}$ (see (2.1)) we then get the claimed formula.

3. Examples. We now analyze some of the consequences of Theorem 2.1, visiting the scenarios A, B, C, and D in section 2.

3.1. Solving univariate polynomials. Let $\mathcal{H}_d[X,Y] \equiv \mathbb{C}^{d+1}$ be the space of degree d homogeneous bivariate polynomials

$$p(X,Y) = \sum_{j=0}^{d} a_j X^j Y^{d-j}, \quad a_0, \ldots, a_d \in \mathbb{C}$$

with the (canonical) Hermitian product given by $\|p\|^2 = \sum_{j=0}^{d} |a_j|^2$.

Let

$$\mathcal{I} = \mathcal{H}_d[X,Y], \quad \mathcal{O} = \mathbb{P}(\mathbb{C}^2), \quad \mathcal{F}(p,z) = p(z),$$

where $z = (X,Y) \in \mathbb{C}^2$. Then a pair (p,z) is in the associated solution variety $\mathcal{V}$ if and only if z is a (projective) zero of p. Setting $Y = 1$ this yields the problem of solving univariate polynomials, although we prefer to use the homogenized version since our result applies directly: the hypotheses of Theorem 2.1 are trivially satisfied. We then have

$$(3.1) \qquad \mathrm{E}_{p \in \mathcal{H}_d[X,Y]} \left(\frac{1}{d} \sum_{z \in \mathbb{P}(\mathbb{C}^2): p(z)=0} \mu(p,z)^2 \right) = \frac{d}{d} = 1,$$

independently of the degree. The condition number (2.5), in this case, has the form

$$(3.2) \qquad \mu(p, (x, 1)) = \frac{\|p\|\|(1, x, \ldots, x^d)\|}{|\hat{p}'(x)|(1 + |x|^2)},$$

where $(x, 1) \in \mathbb{P}(\mathbb{C}^2)$ is a root of p and $\hat{p}(X) = p(X, 1)$ is this univariate polynomial. A random polynomial is obtained by choosing coefficients a_j, associated to the monomial $X^j Y^{d-j}$ (with $0 \leq j \leq d$), independent with distribution $\mathcal{N}_{\mathbb{C}}(0, 1)$. Equation (3.1) then implies that the condition number for such a random polynomial is in general quite small.

3.2. A second view to polynomial solving. Remarkably, as we will now see, the same result on the expected value of the squared condition number obtained in section 3.1 holds if we endow $\mathcal{H}_d[X, Y]$ with the Bombieri–Weyl Hermitian product which makes monomials of different degrees orthogonal and

$$\|X^j Y^{d-j}\|_{BW} = \binom{d}{j}^{-1/2}, \quad (j = 0, \ldots, d)$$

(cf. Remark 2.2). In that classical case the condition number has a very simple expression in terms of the non-homogeneous polynomial $\hat{p}(X) = p(X, 1) = \sum_{j=0}^{d} a_j X^j$ (with $X \in \mathbb{C}$):

$$(3.3) \qquad \mu_{BW}(p, (x, 1)) = \frac{\|\hat{p}\|_{BW} (1 + |x|^2)^{\frac{d}{2}-1}}{|\hat{p}'(x)|},$$

where $\|\hat{p}\|_{BW}$ is the Bombieri–Weyl norm of the polynomial $\hat{p}$, given by

$$\|\hat{p}\|_{BW}^2 = \sum_{j=0}^{d} \binom{d}{j}^{-1} |a_j|^2.$$

See [6] for a proof of (3.3).

This Bombieri–Weyl metric is the standard choice in many papers regarding polynomial system solving after Shub and Smale's work [27]; see also [11, 12].

We can describe this case in our framework taking the input set $\mathcal{I} = \mathcal{H}_d[X, Y]$ (again, bivariate homogeneous polynomials of degree d) but now using the Bombieri–Weyl Hermitian structure. The general definition of the condition number given (2.5) then gives precisely the formula in (3.3). Our Theorem 2.1 then implies

$$(3.4) \qquad \mathrm{E}_{p \in \mathcal{H}_d[X,Y]} \left(\frac{1}{d} \sum_{z \in \mathbb{P}(\mathbb{C}^2): \, p(z)=0} \mu_{BW}(p, z)^2 \right) = \frac{d}{d} = 1,$$

where under the identification $\mathcal{H}_d[X, Y] = \mathbb{C}^{d+1}$, $p(X, Y) = \sum_{j=0}^{d} a_j X^j Y^{d-j}$ is drawn from $a_0, \ldots, a_d$ independent with distribution $\mathcal{N}_{\mathbb{C}}\left(0, \binom{d}{j}\right)$, respectively.

Our claim in (3.4) agrees with the value computed in [8, Theorem 23], where $\sqrt{d}\,\mu_{BW}$ is denoted by μ_{norm}, although the proof is different.

3.3. Lacunary polynomial solving. Now, let $i = \{0, i_1, i_2, \ldots, i_k, d\} \subseteq \{0, \ldots, d\}$ (we identify the case $k = 0$ with $i = \{0, d\}$) be a ordered set of indices and let

$$\mathcal{H}_d^i[X, Y] = \{a_0 Y^d + a_{i_1} X^{i_1} Y^{d-i_1} + \cdots + a_{i_k} X^{i_k} Y^{d-i_k} + a_d X^d : a_0, a_{i_1}, \ldots, a_{i_k}, a_d \in \mathbb{C}\}$$

be the space of polynomials containing only monomials of those degrees. Let

$$\mathcal{I} = \mathcal{H}_d^i[X, Y], \quad \mathcal{O} = \mathbb{P}(\mathbb{C}^2), \quad \mathcal{F}(p, z) = p(z).$$

The hypotheses of Theorem 2.1 are again satisfied (note that the polynomial $X^d - Y^d$ has exactly d projective zeros).

As a subspace of $\mathcal{H}_d[X, Y]$, the input space $\mathcal{H}_d^i[X, Y]$ can be endowed with any of the induced Hermitian products given in the examples above. Then, from Theorem 2.1 we have

$$(3.5) \qquad \mathrm{E}_{p \in \mathcal{H}_d^i[X,Y]} \left(\frac{1}{d} \sum_{z \in \mathbb{P}(\mathbb{C}^2) : \, p(z)=0} \mu(p, z)^2 \right) = \frac{k+1}{d}.$$

This equality describes quantitatively how solving lacunary systems exhibits better stability properties than solving dense polynomials, which we could have qualitatively guessed since the perturbations on the input are more restricted. Note that μ in (3.5) is the condition number adapted to the structure of the input and thus does not obey (3.2) or (3.3).

3.4. Generalized eigenvalue problem. As pointed out above, with

$$\mathcal{I} = \mathbb{C}^{n \times n} \times \mathbb{C}^{n \times n}, \quad \mathcal{O} = \mathbb{P}(\mathbb{C}^2), \quad \mathcal{F}((A, B), (\alpha, \beta)) = \det(\beta A - \alpha B),$$

we get the GEVP. Let us now apply our Theorem 2.1 for the canonical Hermitian inner product on $\mathcal{I}$ which gives the standard Frobenius norm. Note that the degrees of $\mathcal{F}$ in its two entries are $r = s = n$. We check the following.

- The set of eigenvalues is different from the total set $\mathbb{P}(\mathbb{C}^2)$ as long as A and B are not simultaneously singular, that is, in the complement of a $2n^2 - 2$ dimensional subvariety, and there exists (A, B) such that the number of eigenvalues is equal to n.
- For all $(\alpha, \beta) \in \mathbb{P}(\mathbb{C}^2)$, the set $\mathcal{V}_{(\alpha, \beta)}$ of (A, B) such that $\det(\beta A - \alpha B) = 0$ is a $2n^2 - 1$ dimensional variety. Moreover, $\mathcal{V}_{(1,0)}$ is the set of (A, B) such that $\det(B) = 0$ which has degree n.

Thus, the hypotheses of Theorem 2.1 are satisfied. We then have

$$\mathrm{E}_{(A,B) \in \mathbb{C}^{n \times n} \times \mathbb{C}^{n \times n}} \left(\frac{1}{n} \sum_{(\alpha, \beta) \in \mathrm{Eig}(A, B)} \mu((A, B), (\alpha, \beta))^2 \right) = \frac{(2n^2 - 1)n}{n} = 2n^2 - 1.$$

The expected value of the squared condition number for the GEVP is thus essentially equal to the size of the input.

3.5. Polynomial eigenvalue problem. This case is just

$$\mathcal{I} = (\mathbb{C}^{n \times n})^{d+1}, \quad \mathcal{O} = \mathbb{P}(\mathbb{C}^2), \quad \mathcal{F}(A, (\alpha, \beta)) = \det \left(\sum_{j=0}^{d} \alpha^j \beta^{d-j} A_j \right),$$

where $A = (A_0, \ldots, A_d) \in \mathcal{I}$, and we endow $\mathcal{I}$ with the canonical Hermitian inner product. Again the hypotheses of Theorem 2.1 are easily checked with the same tools as in Section 3.4, and we conclude

$$\mathrm{E}_{A \in (\mathbb{C}^{n \times n})^{d+1}} \left(\frac{1}{dn} \sum_{(\alpha, \beta) \in \mathrm{Eig}(A)} \mu(A, (\alpha, \beta))^2 \right) = \frac{((d+1)n^2 - 1)n}{dn} = \frac{(d+1)n^2 - 1}{d},$$

which proves our Theorem 1.1.

3.6. Sparse polynomial eigenvalue problem. The same preceding argument can be applied to sparse versions of the PEVP case. Instead of writing down a generic result, we show how to use it in a particular case: assume for example that we deal with the quadratic eigenvalue problem

$$\mathcal{F}((A,B,C),(\alpha,\beta)) = \det(\alpha^2 A + \alpha\beta B + \beta^2 C) = 0,$$

where we impose on the input matrices some structure. For example, assume that A is diagonal and C is upper triangular.

We check the hypotheses of Theorem 2.1.

- The set of eigenvalues is different from the total set as long as A and C are not simultaneously singular, thus in the complement of a codimension 2 variety, and for input $A = \mathrm{Id}_n$ (the identity matrix), $B = 0$ and generic C the number of solutions is equal to $2n$.
- $\mathcal{V}_{(\alpha,\beta)}$ is a codimension 1 variety for all $(\alpha,\beta) \in \mathcal{O}$ and $\mathcal{V}_{(1,0)}$ is the set of (A,B,C) such that $\det(A) = 0$, that is, a degree n variety.

With the Hermitian structure given by Frobenius product, Theorem 2.1 can thus be applied, and we conclude that in this sparse case there are in general $2n$ eigenvalues and the expected condition number squared equals

$$\mathrm{E}_{A,B,C}\left(\frac{1}{2n}\sum_{(\alpha,\beta)\in\mathrm{Eig}(A,B,C)}\mu((A,B,C),(\alpha,\beta))^2\right) =$$

$$\frac{(n+n^2+\frac{n(n+1)}{2}-1)n}{2n} = \frac{3n^2+3n-2}{4},$$

where the nonzero entries of A, B, and C follow again an independent distribution $\mathcal{N}_{\mathbb{C}}(0,1)$. Note that, as in the case of lacunary polynomials, the condition number μ in this section is the structured condition number and thus does not satisfy (1.2).

3.7. More elaborate eigenvalue problems. We finally include one example of use in the case that $\mathcal{O} \neq \mathbb{P}(\mathbb{C}^2)$. Assume that our problem is the following: on input A, B, C (three complex $n \times n$ matrices), find $(\alpha,\beta,\gamma) \in \mathbb{P}(\mathbb{C}^3)$ such that

$$\alpha\beta + \alpha\gamma + \beta\gamma = 0, \quad \det(\alpha A + \beta B + \gamma C) = 0.$$

This is, in some sense, a system of two homogeneous equations and three variables, thus one would expect the solution set to be a finite collection of points $(\alpha,\beta,\gamma) \in \mathbb{P}(\mathbb{C}^3)$. We can treat this problem inside our framework as follows. Let

$$\mathcal{O} = \{(\alpha,\beta,\gamma) \in \mathbb{P}(\mathbb{C}^3) : \alpha\beta + \alpha\gamma + \beta\gamma = 0\},$$

which is a degree 2 irreducible algebraic subvariety of $\mathbb{P}(\mathbb{C}^3)$. Consider the input space $(\mathbb{C}^{n\times n})^3$ endowed with the Frobenius Hermitian product. We check that the hypotheses of Theorem 2.1 hold.

- If A is nonsingular, then $(1,0,0) \in \mathcal{O}$ is not a solution of

$$\mathcal{F}((A,B,C),(\alpha,\beta,\gamma)) = \det(\alpha A + \beta B + \gamma C) = 0.$$

Similarly, if C is nonsingular then $(0,0,1)$ is not a solution. We conclude that in the complement of a codimension 2 subvariety the number of solutions is

finite. Moreover, for $(A, B, C) = (\mathrm{Id}_n, B, -B)$ where B is diagonal with entries $1, 2, \ldots, n$ the solutions are those points of the form

$$(-m, \gamma + 1, \gamma) \in \mathcal{O}, \quad 1 \leq m \leq n, \quad \gamma \in \mathbb{C}.$$

For these points to be in $\mathcal{O}$ we need

$$-m(2\gamma + 1) + (\gamma + 1)\gamma = 0;$$

that is, for each m there are exactly 2 values of γ that satisfy $(-m, \gamma + 1, \gamma) \in \mathcal{O}$. We thus have found an input with exactly $2n$ solutions.

- $\mathcal{V}_{(\alpha, \beta, \gamma)}$ is a codimension 1 variety for all $(\alpha, \beta, \gamma) \in \mathcal{O}$ and $\mathcal{V}_{(1,0,0)}$ is the set of (A, B, C) such that $\det(A) = 0$, that is, a degree n variety.

We can thus apply Theorem 2.1, and we conclude that there are exactly $2n$ solutions for generic A, B, C and that the expected condition number squared for the problem in this section equals

$$\mathrm{E}_{A,B,C} \left(\frac{1}{2n} \sum_{(\alpha,\beta,\gamma) \in \mathcal{O} : \mathcal{F}((A,B,C),(\alpha,\beta,\gamma))=0} \mu((A, B, C), (\alpha, \beta, \gamma))^2 \right) =$$

$$\frac{(3n^2 - 1)n}{n} = 3n^2 - 1,$$

where A, B, C have random independent entries with distribution $\mathcal{N}_{\mathbb{C}}(0, 1)$.

Note that the choice of the variety $\alpha\beta + \alpha\gamma + \beta\gamma$ in this example is not crucial: for a generic homogeneous equation defined in $\mathbb{P}(\mathbb{C}^3)$, the solution set is an irreducible algebraic variety and in many cases one can very easily check that the hypotheses of our theorem still hold.

Appendix A. A useful integral. In this section we prove the following lemma.

LEMMA A.1. *Let $\mathcal{J} \subseteq \mathbb{C}^a$ be a homogeneous complex algebraic variety of degree d and dimension n. Then,*

$$\int_{p \in \mathcal{J}} \|p\|^2 e^{-\|p\|^2} \, dp = \pi^n \, nd.$$

Proof. Let $\Pi : \mathcal{J} \to \mathbb{P}(\mathcal{J})$ be the restriction to $\mathcal{J}$ of the quotient map $\mathbb{C}^a \to \mathbb{P}(\mathbb{C}^a)$. Its Normal Jacobian is equal to

$$NJ(\Pi)(p) = \frac{1}{\|p\|^{2n-2}}.$$

(In order to compute this Normal Jacobian just consider any orthonormal basis of $T_p\mathcal{J}$ whose last vector is $\dot{p} = p$ and use the definition of the Normal Jacobian given in Appendix B). The coarea formula (Theorem B.1 below) applied to π then yields

$$\int_{p \in \mathcal{J}} \|p\|^2 e^{-\|p\|^2} \, dp = \int_{q \in \mathbb{P}(\mathcal{J})} \int_{p \in \mathcal{J} : \pi(p) = q} \|p\|^{2n-2+2} e^{-\|p\|^2} \, dp \, dq =$$

$$Vol(\mathbb{P}(\mathcal{J})) \int_{w \in \mathbb{C}} |w|^{2n} e^{-|w|^2} \, dw.$$

Since $\mathbb{P}(\mathcal{J})$ has degree d and dimension $n-1$, from (2.1) its volume in the projective space satisfies

$$Vol(\mathbb{P}(\mathcal{J})) = Vol(\mathbb{P}(\mathbb{C}^n))\,d = \frac{\pi^{n-1}}{\Gamma(n)}\,d.$$

On the other hand, using polar coordinates we have

$$\int_{w\in\mathbb{C}} |w|^{2n} e^{-|w|^2}\,dw = 2\pi \int_0^\infty t^{2n+1} e^{-t^2}\,dt = \pi \int_0^\infty s^n e^{-s}\,ds = \pi\Gamma(n+1).$$

The lemma follows. □

Appendix B. The coarea formula. For completeness of this manuscript we include here the smooth coarea formula, a version for Riemannian manifolds of a well known generalization of the change of variables theorem and Fubini's theorem discovered by Federer. Here we follow [11, page 241] or [12, Theorem 17.8]. First we define the Normal Jacobian of a smooth mapping $\varphi : M \to N$ where M and N are smooth manifolds of dimensions $m \geq n$: for $x \in M$, let $v_1, \ldots, v_n, v_{n+1}, \ldots, v_m$ be an orthonormal basis of the tangent space $T_x M$, with $D\varphi(x)v_{n+1} = \cdots = D\varphi(x)v_m = 0$. Then, $NJ\varphi(x)$ is the volume of the parallelepiped spanned by $D\varphi(x)v_1, \ldots, D\varphi(x)v_n$ in $T_{\varphi(x)}N$. In other words, $NJ\varphi(x)$ is the absolute value of the determinant of the derivative $D\varphi(x)$ restricted to the orthogonal complement of its kernel. If $D\varphi$ is not surjective, the Normal Jacobian is set to 0.

THEOREM B.1 (the smooth coarea formula). *Let M, N be Riemannian manifolds of respective dimensions $m \geq n$, and let $\varphi : M \to N$ be a smooth surjective surjection (i.e., φ is surjective and the derivative $D\varphi$ is surjective at every point of M). Then, for any positive measurable function $\phi : M \to [0, \infty)$ we have*

$$\int_{x\in M} \phi(x)\,dx = \int_{y\in N} \int_{x\in\varphi^{-1}(y)} \frac{\phi(x)}{NJ\varphi(x)}\,dx\,dy$$

and

$$\int_{x\in M} NJ\varphi(x)\phi(x)\,dx = \int_{y\in N} \int_{x\in\varphi^{-1}(y)} \phi(x)\,dx\,dy$$

(in particular, the left-hand side is finite if and only if so is the right hand side in both equalities).

Note that in Theorem B.1 the inner integrals are to be understood as integrals in the submanifolds $\varphi^{-1}(y)$ with the metric inherited from M. These sets are submanifolds for a.e. $y \in N$ by Sard's theorem.

REFERENCES

[1] D. ARMENTANO, *Stochastic perturbations and smooth condition numbers*, J. Complexity, 26 (2010), pp. 161–171, https://doi.org/10.1016/j.jco.2010.01.003.
[2] D. ARMENTANO, C. BELTRÁN, P. BÜRGISSER, F. CUCKER, AND M. SHUB, *A stable, polynomial-time algorithm for the eigenpair problem*, J. Eur. Math. Soc., 20 (2018), pp. 1375–1437, https://doi.org/10.4171/JEMS/789.
[3] D. ARMENTANO AND F. CUCKER, *A randomized homotopy for the Hermitian eigenpair problem*, Found. Comput. Math., 15 (2015), pp. 281–312, https://doi.org/10.1007/s10208-014-9217-9.

[4] J. L. AURENTZ, R. VANDEBRIL, AND D. S. WATKINS, *Fast computation of the zeros of a polynomial via factorization of the companion matrix*, SIAM J. Sci. Comput., 35 (2013), pp. A255–A269, https://doi.org/10.1137/120865392.

[5] Z. BAI AND Y. SU, *SOAR: A second-order Arnoldi method for the solution of the quadratic eigenvalue problem*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 640–659, https://doi.org/10.1137/S0895479803438523.

[6] C. BELTRÁN, *A facility location formulation for stable polynomials and elliptic Fekete points*, Found. Comput. Math., 15 (2015), pp. 125–157, https://doi.org/10.1007/s10208-014-9213-0.

[7] C. BELTRÁN AND K. KOZHASOV, *The real polynomial eigenvalue problem is well conditioned on the average*, preprint, arXiv:1802.07493, https://arxiv.org/abs/1802.07493.

[8] C. BELTRÁN AND L. M. PARDO, *Fast linear homotopy to find approximate zeros of polynomial systems*, Found. Comput. Math., 11 (2011), pp. 95–129, https://doi.org/10.1007/s10208-010-9078-9.

[9] D. A. BINI, P. BOITO, Y. EIDELMAN, L. GEMIGNANI, AND I. GOHBERG, *A fast implicit QR eigenvalue algorithm for companion matrices*, Linear Algebra Appl., 432 (2010), pp. 2006–2031, https://doi.org/10.1016/j.laa.2009.08.003.

[10] D. A. BINI, L. GEMIGNANI, AND V. Y. PAN, *Fast and stable QR eigenvalue algorithms for generalized companion matrices and secular equations*, Numer. Math., 100 (2005), pp. 373–408, https://doi.org/10.1007/s00211-005-0595-4.

[11] L. BLUM, F. CUCKER, M. SHUB, AND S. SMALE, *Complexity and real computation*, Springer-Verlag, New York, 1998, https://doi.org/10.1007/978-1-4612-0701-6.

[12] P. BÜRGISSER AND F. CUCKER, *Condition*, Grundlehren der Mathematischen Wissenschaften 349 [Fundamental Principles of Mathematical Sciences], Springer, Heidelberg, 2013, https://doi.org/10.1007/978-3-642-38896-5.

[13] P. BÜRGISSER, F. CUCKER, AND E. ROCHA CARDOZO, *On the condition of the zeros of characteristic polynomials*, J. Complexity, 42 (2017), pp. 72–84, https://doi.org/10.1016/j.jco.2017.03.004.

[14] Z. CHEN AND J. J. DONGARRA, *Condition numbers of Gaussian random matrices*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 603–620, https://doi.org/10.1137/040616413.

[15] F. DE TERÁN, F. M. DOPICO, AND J. PÉREZ, *Backward stability of polynomial root-finding using Fiedler companion matrices*, IMA J. Numer. Anal., 36 (2016), pp. 133–173, https://doi.org/10.1093/imanum/dru057.

[16] J.-P. DEDIEU AND F. TISSEUR, *Perturbation theory for homogeneous polynomial eigenvalue problems*, Linear Algebra Appl., 358 (2003), pp. 71–94, https://doi.org/10.1016/S0024-3795(01)00423-2.

[17] J. W. DEMMEL, *On condition numbers and the distance to the nearest ill-posed problem*, Numer. Math., 51 (1987), pp. 251–289, https://doi.org/10.1007/BF01400115.

[18] A. EDELMAN, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 543–560, https://doi.org/10.1137/0609045.

[19] L. GRAMMONT, N. J. HIGHAM, AND F. TISSEUR, *A framework for analyzing nonlinear eigenproblems and parametrized linear systems*, Linear Algebra Appl., 435 (2011), pp. 623–640, https://doi.org/10.1016/j.laa.2009.12.038.

[20] S. HAMMARLING, C. J. MUNRO, AND F. TISSEUR, *An algorithm for the complete solution of quadratic eigenvalue problems*, ACM Trans. Math. Software, 39 (2013), https://doi.org/10.1145/2450153.2450156.

[21] N. J. HIGHAM, R.-C. LI, AND F. TISSEUR, *Backward error of polynomial eigenproblems solved by linearization*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1218–1241, https://doi.org/10.1137/060663738.

[22] N. J. HIGHAM, D. S. MACKEY, AND F. TISSEUR, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1005–1028, https://doi.org/10.1137/050628283.

[23] S. A. ŁOJASIEWICZ, *Introduction to complex analytic geometry*, Birkhäuser Verlag, Basel, 1991, https://doi.org/10.1007/978-3-0348-7617-9. Translated from the Polish by Maciej Klimek.

[24] P. LANCASTER AND I. ZABALLA, *On the inverse symmetric quadratic eigenvalue problem*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 254–278, https://doi.org/10.1137/130905216.

[25] K. MEERBERGEN, *The quadratic Arnoldi method for the solution of the quadratic eigenvalue problem*, SIAM J. Matrix Anal. Appl., 30 (2008/09), pp. 1463–1482, https://doi.org/10.1137/07069273X.

[26] V. MEHRMANN AND H. VOSS, *Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods*, GAMM Mitt. Ges. Angew. Math. Mech., 27 (2004), pp. 121–152 (2005), https://doi.org/10.1002/gamm.201490007.

[27] M. Shub and S. Smale, *Complexity of Bézout's theorem.* I. *Geometric aspects*, J. Amer. Math. Soc., 6 (1993), pp. 459–501, https://doi.org/10.2307/2152805.

[28] F. Tisseur, *Backward error and condition of polynomial eigenvalue problems*, in Proceedings of the International Workshop on Accurate Solution of Eigenvalue Problems (University Park, PA, 1998), vol. 309, 2000, pp. 339–361, https://doi.org/10.1016/S0024-3795(99)00063-4.

[29] L. N. Trefethen and D. Bau, III, *Numerical linear algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997, https://doi.org/10.1137/1.9780898719574.

[30] M. Van Barel, R. Vandebril, P. Van Dooren, and K. Frederix, *Implicit double shift QR-algorithm for companion matrices*, Numer. Math., 116 (2010), pp. 177–212, https://doi.org/10.1007/s00211-010-0302-y.