J. Symbolic Comput., vol. 75, pp. 209–218, July/August 2016.
Submitted Dec. 8, 2014, accepted August 18, 2015.

Sparse Multivariate Function Recovery With a Small Number of Evaluations*

Erich L. Kaltofen[1] and Zhengfeng Yang[2]

[1]Department of Mathematics, North Carolina State University,
Raleigh, North Carolina 27695-8205, USA
kaltofen@math.ncsu.edu; http://www.math.ncsu.edu/~kaltofen

[2]Shanghai Key Laboratory of Trustworthy Computing, SEI
East China Normal University, Shanghai 200062, China
zfyang@sei.ecnu.edu.cn

Abstract

In [Kaltofen and Yang, Proc. ISSAC 2014] we give an algorithm based algebraic error-correcting decoding for multivariate sparse rational function interpolation from evaluations that can be numerically inaccurate and where several evaluations can have severe errors ("outliers"). Our 2014 algorithm can interpolate a sparse multivariate rational function from evaluations where the error rate is $1/q$ is quite high, say $q = 5$.

For the algorithm with exact arithmetic and exact values at non-erroneous points, one avoids quadratic oversampling by using random evaluation points. Here we give the full probabilistic analysis for this fact, thus providing the missing proof to Theorem 2.1 in Section 2 of our ISSAC 2014 paper. Our argumentation already applies to our original 2007 sparse rational function interpolation algorithm [Kaltofen, Yang and Zhi, Proc. SNC 2007], where we have experimentally observed that for T unknown non-zero coefficients in a sparse candidate ansatz one only needs $T + O(1)$ evaluations rather than $O(T^2)$ (cf. Candès and Tao sparse sensing), the latter of which we have proved in 2007. Here we prove that $T + O(1)$ evaluations at random points indeed suffice.

1. Our Vector-of-Functions Recovery Setting

We now present the setting of our theorem on the required number of samples for rational function recovery. For the full background including the references to the extant literature and our error-tolerant multivariate rational function interpolation algorithm, its implementation and observed experimental data we refer to our paper (Kaltofen and Yang 2014).

*This research was supported in part by the National Science Foundation under Grant CCF-1115772 (Kaltofen).

We interpolate *a vector* of multivariate *sparse* rational functions with a common denominator:

$$\left[\frac{f^{\langle 1\rangle}}{g},\ldots,\frac{f^{\langle s\rangle}}{g}\right]\in\mathsf{K}(x_1,\ldots,x_n)^s,\quad g\neq 0. \tag{1}$$

Note that the fractions $f^{\langle\sigma\rangle}/g$ are not necessarily reduced, and that may even have $\mathrm{GCD}(g,\mathrm{GCD}_\sigma(f^{\langle\sigma\rangle}))\neq 1$, because reduction by GCD can affect the sparsity of the fraction, such as $(x_1^d-x_2^d)/(x_1-x_2)$. We assume that we have for all σ, $1\leq\sigma\leq s$, sets of terms $D_f^{\langle\sigma\rangle}\supseteq\mathrm{supp}(f^{\langle\sigma\rangle})$ that constitute maximal sparse supports, and a maximal sparse support set $D_g\supseteq\mathrm{supp}(g)$ for the terms in the common denominator g. See the Appendix (Section 3 below) for a definition of the support and the meaning of all used symbols. Our algorithms (Kaltofen et al. 2007; Kaltofen and Yang 2013, 2014) follow the variable-by-variable process by Zippel (Zippel 1979), which yield those sparse support supersets in each iteration. We suppose that we can evaluate the vector (1) ("probe the black box") at values for the variables, $(x_1,\ldots,x_n)\leftarrow(\xi_{1,\ell},\ldots,\xi_{n,\ell})\in\mathsf{K}^n$, for all L evaluations $0\leq\ell\leq L-1$, where the $\xi_{\mu,\ell}$ are chosen in a certain way, e.g., selected randomly and uniformly from a finite subset $S\subseteq\mathsf{K}$. As in (Kaltofen and Yang 2013), the obtained vector $[\beta_\ell^{\langle 1\rangle},\ldots,\beta_\ell^{\langle s\rangle}]\in(\mathsf{K}\cup\{\infty\})^s$ can be incorrect in one or more components for $k\leq E$ evaluations $\ell=\lambda_1,\ldots,\lambda_k$, that is

$$\forall\kappa,1\leq\kappa\leq k\colon\exists\sigma,1\leq\sigma\leq s\colon\ \frac{f^{\langle\sigma\rangle}}{g}(\xi_{1,\lambda_\kappa},\ldots\xi_{n,\lambda_\kappa})\neq\beta_{\lambda_\kappa}^{\langle\sigma\rangle}, \tag{2}$$

$$\forall\ell\notin\{\lambda_1,\ldots,\lambda_k\}\colon\forall\sigma,1\leq\sigma\leq s\colon\ \frac{f^{\langle\sigma\rangle}}{g}(\xi_{1,\ell},\ldots\xi_{n,\ell})=\beta_\ell^{\langle\sigma\rangle}. \tag{3}$$

Here E is predetermined, for instance from the error rate (Kaltofen and Yang 2014, Remark 1.1), and the locations of the errors are unknown. As in (Kaltofen and Yang 2013) we set *all* components of a vector $=\infty$ if $g(\xi_{1,\ell},\ldots,\xi_{n,\ell})=0$, that even for those components with $f^{\langle\sigma\rangle}(\xi_{1,\ell},\ldots,\xi_{n,\ell})=0$, but false vectors full of ∞'s can appear for $g(\xi_{1,\lambda_\kappa},\ldots,\xi_{n,\lambda_\kappa})\neq 0$. We can identify vectors that contain both ∞ and a field element as erroneous. Errors are dealt with by interpolating $(f^{\langle\sigma\rangle}\Lambda)/(g\Lambda)$ à la (Kaltofen and Pernet 2013; Kaltofen and Yang 2013) where $\Lambda=(x_{n_1}-\xi_{n_1,\lambda_1})\cdots(x_{n_1}-\xi_{n_1,\lambda_k})$ is an error locator polynomial for a chosen n_1 with $1\leq n_1\leq n$. We have the maximal supports

$$\left.\begin{aligned}D_{f,E;n_1}^{\langle\sigma\rangle}&=\{\tau x_{n_1}^\nu\mid\tau\in D_f^{\langle\sigma\rangle},0\leq\nu\leq E\}\supseteq\mathrm{supp}(f^{\langle\sigma\rangle}\Lambda),\\ D_{g,E;n_1}&=\{\tau x_{n_1}^\nu\mid\tau\in D_g,0\leq\nu\leq E\}\ \ \supseteq\mathrm{supp}(g\Lambda).\end{aligned}\right\} \tag{4}$$

Now we limit the sparse supports of polynomials with unknown coefficients $\Phi^{\langle\sigma\rangle}$ and Ψ to the term sets (4). From (2) and (3) we obtain linear homogeneous equations for the coefficients of $\Phi^{\langle\sigma\rangle}$, Ψ:

$$\left.\begin{aligned}&\Phi^{\langle\sigma\rangle}(\xi_{1,\ell},\ldots,\xi_{n,\ell})-\beta_\ell^{\langle\sigma\rangle}\Psi(\xi_{1,\ell},\ldots,\xi_{n,\ell})=0,\\ &\qquad\text{for } 0\leq\ell\leq L-1,1\leq\sigma\leq s \text{ with } \beta_\ell^{\langle\sigma\rangle}\neq\infty,\\ &\Psi(\xi_{1,\ell},\ldots,\xi_{n,\ell})=0,\\ &\qquad\text{for } 0\leq\ell\leq L-1 \text{ with } \beta_\ell^{\langle 1\rangle}=\cdots=\beta_\ell^{\langle s\rangle}=\infty,\\ &\text{with } \mathrm{supp}(\Phi^{\langle\sigma\rangle})\subseteq D_{f,E;n_1}^{\langle\sigma\rangle} \text{ for } 1\leq\sigma\leq s,\ \mathrm{supp}(\Psi)\subseteq D_{g,E;n_1}.\end{aligned}\right\} \tag{5}$$

Note that $\Phi^{\langle\sigma\rangle} \leftarrow f^{\langle\sigma\rangle}\Lambda$, $\Psi \leftarrow g\Lambda$ solve (5). We call any solution $(\Phi^{\langle 1\rangle}, \ldots, \Phi^{\langle s\rangle}, \Psi)$ of (5) an interpolant. We seek a (minimal) L and $\xi_{\mu,\ell}$ such that *all* solutions of (5) satisfy

$$\forall \sigma, 1 \le \sigma \le s\colon \Phi^{\langle\sigma\rangle} g = f^{\langle\sigma\rangle}\Psi, \quad \text{with } \operatorname{supp}(\Phi^{\langle\sigma\rangle}) \subseteq D^{\langle\sigma\rangle}_{f,E;n_1},\ \operatorname{supp}(\Psi) \subseteq D_{g,E;n_1}. \tag{6}$$

We call (6) the Welch-Berlekamp property. Then any non-zero solution vector to (6) satisfies

$$\left[\frac{\Phi^{\langle 1\rangle}}{\Psi}, \ldots, \frac{\Phi^{\langle s\rangle}}{\Psi}\right] = \left[\frac{f^{\langle 1\rangle}}{g}, \ldots, \frac{f^{\langle s\rangle}}{g}\right].$$

The equation $\Phi^{\langle\sigma\rangle} g = f^{\langle\sigma\rangle}\Psi$ is the equivalence test for the field of quotients construction for $\mathsf{K}[x_1, \ldots, x_n]$, which works for any integral domain and does not require a greatest common divisor operation. That is the underpinning reason why the sparse fractions in (1) can be left unreduced.

The key theorem below states that at random points the Welch-Berlekamp property is achieved from almost square systems. The following Theorem 1.1 is Theorem 2.1 in (Kaltofen and Yang 2014).

Theorem 1.1. *Let*

$$L = |D_{g,E;n_1}| + (\max_{1\le\sigma\le s} |D^{\langle\sigma\rangle}_{f,E;n_1}|) - 1, \quad M^{\langle\sigma\rangle} = |D_{g,E;n_1}| + |D^{\langle\sigma\rangle}_{f,E;n_1}|, \tag{7}$$

and let all $\xi_{\mu,\ell}$, where $1 \le \mu \le n$ and $0 \le \ell \le L-1$, be randomly and uniformly selected from a finite subset $S \subseteq \mathsf{K}$. Then the probability that all $(s+1)$-tuples $(\Phi^{\langle 1\rangle}, \ldots, \Phi^{\langle s\rangle}, \Psi)$ that are interpolants to (5) satisfy the Welch-Berlekamp property (6) is bounded from below as

$$\ge 1 - \frac{1}{|S|}\sum_{\sigma=1}^{s} (M^{\langle\sigma\rangle} - E - 1)\,(\max\{\deg(\tau_f) \mid \tau_f \in D^{\langle\sigma\rangle}_f\} + \max\{\deg(\tau_g) \mid \tau_g \in D_g\} + E).$$

Theorem 1.1 as stated is a five-fold generalization to Cauchy interpolation: multivariate instead of univariate (μ and n in Theorem 1.1), sparse instead of dense (the D's in Theorem 1.1), vector instead of a single function (σ and s in Theorem 1.1), error-correction instead of always correct values (E and n_1 in Theorem 1.1), values at poles indicated by ∞ instead of disallowed (equation for Ψ in (5)).

Before giving our proof to the above Theorem 1.1, we demonstrate our technique on a simpler problem, the non-singularity of sparse Fourier matrices.

Lemma 1.2. *Let $0 \le d_1 < d_2 < \cdots < d_t < \bar{d}$ be integers, and let $\eta_1, \ldots, \eta_t$ be uniformly randomly selected integers with $0 \le \eta_\ell \le \bar{d}-1$ for all $1 \le \ell \le t$. Then the matrix $[e^{2\pi \boldsymbol{i} d_j \eta_\ell/\bar{d}}]_{1\le j,\ell\le t}$ is non-singular with probability $\ge 1 - (d_1 + \cdots + d_t)/\bar{d}$.*

Proof. The matrix $A(v_1, \ldots, v_t) = [v_\ell^{d_j}]_{1\le j,\ell\le t} \in \mathbb{C}[v_1, \ldots, v_t]^{t\times t}$ is non-singular because the variable substitution $v_\ell \leftarrow v_1^{\ell-1}$ yields $A(1, v_1, v_1^2, \ldots, v_1^{t-1})$ as a (transposed) Vandermonde matrix and the terms $v_1^{d_1}, \ldots, , v_1^{d_t}$ are distinct. Another argument would be that the term $v_1^{d_1} v_2^{d_2} \cdots v_n^{d_n}$ in the minor expansion of the determinant of $A(v_1, \ldots, v_t)$ does not cancel, but in our proof of Theorem 1.1 we will make use of a similar substitution as the one above. The total degree $\deg(\det(A)) \le d_1 + \cdots + d_t$, so by the DeMillo-Lipton-Schwartz-Zippel Lemma

applied to the sample set $S = \{e^{2\pi i\eta/\bar{d}} \mid 0 \le \eta \le \bar{d}-1\}$ we obtain the stated probability. □

Note that Lemma 1.2 generalizes the well-known non-singularity for prime $\bar{d}$ (see (Tao 2005, Lemma 1.3)) to composite $\bar{d}$ under the assumption of random η_ℓ.

2. Proof of Theorem 1.1

The proof of Theorem 1.1 must account for several complications, two of which are: 1. the vector of fractions (1) is sparse and unreduced, so given degree bounds and term-supersets for the numerators and the denominator on input, one may have several sparse interpolants: for instance x_1/x_2 and $x_1^2/(x_1x_2)$; 2. not only are the error locations λ_κ unknown, the actual number of errors, k, is also unknown. On input, one has a bound $E \ge k$.

We split the proof into 3 auxiliary lemmas followed by the main argument. Our first auxiliary lemma relates solutions with the Welch-Berlekamp property (6) to interpolants of (5); it appeared first for $s = 1$ in (Kaltofen and Yang 2013, *Note added to Remark 2.2,* in the posting on Kaltofen's web site on July 14, 2013).

Lemma 2.1. *For any $L \ge k \ge 0$, any $E \ge 0$ and any evaluations $\xi_{\mu,\ell} \in \mathsf{K}$, where $1 \le \mu \le n$, $0 \le \ell \le L-1$, consider the solution $(s+1)$-tuples $(\Phi^{\langle 1\rangle}, \ldots, \Phi^{\langle s\rangle}, \Psi)$ to the homogeneous linear equations in their coefficients*

$$\forall \sigma, 1 \le \sigma \le s\colon \Phi^{\langle\sigma\rangle} g = f^{\langle\sigma\rangle}\Psi \qquad (\textit{that is, } (6)) \tag{8}$$

$$\Phi^{\langle\sigma\rangle}(\xi_{1,\lambda_\kappa}, \ldots, \xi_{n,\lambda_\kappa}) - \beta_{\lambda_\kappa}^{\langle\sigma\rangle}\Psi(\xi_{1,\lambda_\kappa}, \ldots, \xi_{n,\lambda_\kappa}) = 0,$$
$$\textit{for } 1 \le \kappa \le k, 1 \le \sigma \le s \textit{ with } \beta_{\lambda_\kappa}^{\langle\sigma\rangle} \ne \infty, \tag{9}$$

$$\Psi(\xi_{1,\lambda_\kappa}, \ldots, \xi_{n,\lambda_\kappa}) = 0, \textit{ for } 1 \le \kappa \le k \textit{ with } \beta_{\lambda_\kappa}^{\langle 1\rangle} = \cdots = \beta_{\lambda_\kappa}^{\langle s\rangle} = \infty, \tag{10}$$

$$\Psi(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) = 0,$$
$$\textit{with } \ell \notin \{\lambda_1, \ldots, \lambda_k\} \textit{ and } g(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) = 0, \forall\sigma\colon f^{\langle\sigma\rangle}(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) = 0. \tag{11}$$

$$\textit{with } \operatorname{supp}(\Phi^{\langle\sigma\rangle}) \subseteq D_{f,E;n_1}^{\langle\sigma\rangle} \textit{ for } 1 \le \sigma \le s, \operatorname{supp}(\Psi) \subseteq D_{g,E;n_1}.$$

All those solution tuples must be interpolants of (5).

Proof of Lemma 2.1. For $\ell \notin \{\lambda_1, \ldots, \lambda_k\}$ and $\beta_\ell^{\langle\sigma\rangle} \ne \infty$ we have for all σ:

$$\begin{aligned}\beta_\ell^{\langle\sigma\rangle}(g\ \Psi)(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) &= (f^{\langle\sigma\rangle}\ \Psi)(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) \text{ (by the definition (3))}\\ &= (\Phi^{\langle\sigma\rangle}\ g)(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) \text{ (by (8))}.\end{aligned}$$

Dividing by $g(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) \ne 0$ yields (5) for this case. For $\ell \in \{\lambda_1, \ldots, \lambda_k\}$ and $\beta_\ell^{\langle\sigma\rangle} = \infty$ ("false pole") we have (5) from (10). Finally, for $\ell \notin \{\lambda_1, \ldots, \lambda_k\}$ and $\beta_\ell^{\langle\sigma\rangle} = \infty$ ("true pole") we have $(f^{\langle\sigma\rangle}\ \Psi)(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) = (\Phi^{\langle\sigma\rangle}\ g)(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) = 0$ by (8) and $g(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) = 0$. If one $f^{\langle\sigma\rangle}(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) \ne 0$ we get $\Psi(\xi_{1,\ell}, \ldots, \xi_{n,\ell}) = 0$ of (5). Otherwise we use (11). □

Our second auxiliary lemma gives an upper bound on L so that all interpolants of (5) for certain $\xi_{\mu,\ell}$ are in the described subspace of Lemma 2.1, meaning that they satisfy the Welch-Berlekamp property (6). The argument for $s = 1$ is already in (Kaltofen et al. 2007, Section 4.1).

Lemma 2.2. *Let* $L_\times = |D_{g,E;n_1}| \times (\max_{1\le\sigma\le s} |D^{\langle\sigma\rangle}_{f,E;n_1}|)$ *and let* $\xi_{\mu,\ell} = \xi_\mu^\ell \in \mathsf{K}$, *where* $1 \le \mu \le n$ *and* $0 \le \ell \le L_\times - 1$, *such that for* $D^{\langle\sigma\rangle}_{f,E;n_1} \times D_{g,E;n_1} = \{\tau_f\tau_g \mid \tau_f \in D^{\langle\sigma\rangle}_{f,E;n_1}, \tau_g \in D_{g,E;n_1}\}$ *we have*

$$\tau_1(\xi_1,\ldots,\xi_n) \ne \tau_2(\xi_1,\ldots,\xi_n) \text{ for all } \tau_1,\tau_2 \in D^{\langle\sigma\rangle}_{f,E;n_1} \times D_{g,E;n_1}, \tau_1 \ne \tau_2, \tag{12}$$

(see (Kaltofen and Yang 2013, Assumption 4)). Then all interpolants $(s+1)$*-tuples* $(\Phi^{\langle 1\rangle}, \ldots, \Phi^{\langle s\rangle}, \Psi)$ *of* (5) *satisfy the Welch-Berlekamp property* (6).

Proof of Lemma 2.2. Because for $\beta_\ell^{\langle\sigma\rangle} \ne \infty$ we have

$$(f^{\langle\sigma\rangle}\Lambda)(\xi_1^\ell,\ldots,\xi_n^\ell) = \beta_\ell^{\langle\sigma\rangle}(g\Lambda)(\xi_1^\ell,\ldots,\xi_n^\ell)$$

we get from (5) for $\beta_\ell^{\langle\sigma\rangle} \ne \infty$ that $(\Phi^{\langle\sigma\rangle}g\Lambda - f^{\langle\sigma\rangle}\Lambda\Psi)(\xi_1^\ell,\ldots,\xi_n^\ell) = 0$. For $\beta_\ell^{\langle\sigma\rangle} = \infty$ we have at both true or false poles $(\Lambda g)(\xi_1^\ell,\ldots,\xi_n^\ell) = \Psi(\xi_1^\ell,\ldots,\xi_n^\ell) = 0$, the latter by (5), and hence again $(\Phi^{\langle\sigma\rangle}g\Lambda - f^{\langle\sigma\rangle}\Lambda\Psi)(\xi_1^\ell,\ldots,\xi_n^\ell) = 0$. Thus the coefficient vectors of $\Phi^{\langle\sigma\rangle}g\Lambda - f^{\langle\sigma\rangle}\Lambda\Psi$ are nullspace vectors of the matrix with entries $\tau(\xi_1,\ldots,\xi_n)^\ell$ where $\tau \in D^{\langle\sigma\rangle}_{f,E;n_1} \times D_{g,E;n_1}$. By our assumption (12) the matrix is transposed Vandermonde with distinct entries in each row, and has $L_\times$ rows, which is $\ge$ the number of terms in $\mathrm{supp}(\Phi^{\langle\sigma\rangle}g\Lambda - f^{\langle\sigma\rangle}\Lambda\Psi)$. Therefore for all σ the coefficient vectors of $\Phi^{\langle\sigma\rangle}g\Lambda - f^{\langle\sigma\rangle}\Lambda\Psi$ are zero. □

Remark 2.1. For $n = 1$ and dense support sets $D^{\langle\sigma\rangle}_{f,E;n_1} = \{1, x_1, x_1^2, \ldots\}$, $D_{g,E;n_1} = \{1, x_1, x_1^2, \ldots\}$ we may choose $\xi_{1,\ell} = \hat\xi_\ell \in \mathsf{K}$ with $\hat\xi_{\ell_1} \ne \hat\xi_{\ell_2}$ for all $\ell_1 \ne \ell_2$. Then the coefficient matrix for $(\Phi^{\langle\sigma\rangle}g\Lambda - f^{\langle\sigma\rangle}\Lambda\Psi)(\hat\xi_\ell)$ is a non-zero Vandermonde matrix and the Lemma holds. See also (Olshevsky and Shokrollahi 2003). □

The third lemma is the crucial idea in (Kaltofen and Yang 2013, *Note added to Remark 2.2,* in the posting on Kaltofen's web site on July 14, 2013) that reduces $L_\times$ of Lemma 2.2. We will evaluate the black box for (1) at symbols $v_{\mu,\ell} \in \mathsf{K}(\ldots, v_{\mu,\ell}, \ldots)$. It is not required from the black box to allow such elements (in transcendental extensions of K) as arguments, we solely use it for purpose of proof.

Lemma 2.3. *Let* $L_+ = |D_{g,E;n_1}| + (\max_{1\le\sigma\le s} |D^{\langle\sigma\rangle}_{f,E;n_1}|) - E - 1$. *Suppose* $\xi_{\mu,\ell} = v_{\mu,\ell}$ *is a new symbol (variable), for each* $1 \le \mu \le n$, $0 \le \ell \le L_+ - 1$. *We assume that* $k = 0$, *that is, there are no erroneous evaluations, so that* $\beta_\ell^{\langle\sigma\rangle} = (f^{\langle\sigma\rangle}/g)(v_{1,\ell},\ldots,v_{n,\ell}) \in \mathsf{K}(v_{1,0},\ldots,v_{n,L_+-1})$ *for all* ℓ. *Note that at vectors of* n *distinct variables there cannot be true poles. Then all interpolants* $(s+1)$*-tuples of* (5) *for* $L = L_+$ *over* $\mathsf{K}(v_{1,0},\ldots,v_{n,L_+-1})$ *satisfy the Welch-Berlekamp property* (6).

Proof of Lemma 2.3. If $\xi_{\mu,\ell} = v_\mu^\ell$, where v_μ are symbols for variables (transcendental elements), and $L_\times = |D_{g,E;n_1}| \times (\max_{1\le\sigma\le s} |D^{\langle\sigma\rangle}_{f,E;n_1}|)$ evaluations are used, Lemma 2.2 gives (6) for all interpolants of (5) over $\mathsf{K}(v_1,\ldots,v_n)$. Note that (12) is satisfied for the variables v_μ. We first show the same for $\xi_{\mu,\ell} = v_{\mu,\ell}$. As in the proof of Lemma 2.2 we have $(\Phi^{\langle\sigma\rangle}g - f^{\langle\sigma\rangle}\Psi)(v_{1,\ell},\ldots,v_{n,\ell}) = 0$ for all ℓ, $0 \le \ell \le L_\times - 1$; note that $k = 0 \Rightarrow \Lambda = 1$. Solving for the coefficient vector of $\Phi^{\langle\sigma\rangle}g - f^{\langle\sigma\rangle}\Psi$ over $\mathsf{K}(v_{1,0},\ldots,v_{n,L_\times-1})$, the arising coefficient matrix has entries $\tau(v_{1,\ell},\ldots,v_{n,\ell})$ for $\tau \in D^{\langle\sigma\rangle}_{f,E;n_1} \times D_{g,E;n_1}$. The matrix has full column rank

because it does so when substituting $v_{\mu,\ell} \leftarrow v_\mu^\ell$ (cf. the proof of Lemma 1.2 above), which are evaluations in the form of Lemma 2.2. Therefore the coefficient vector of $\Phi^{\langle\sigma\rangle} g - f^{\langle\sigma\rangle}\Psi$ is zero.

We now can reduce the number of equations in (5) (for $\xi_{\mu,\ell} = v_{\mu,\ell}$) without enlarging the set of interpolants of (5). Fix a σ, that is, the case $s = 1$: we have shown just above that all solution pairs $(\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle})$ to the equations

$$\Phi^{\langle\sigma\rangle}(v_{1,\ell},\ldots,v_{n,\ell}) - \beta_\ell^{\langle\sigma\rangle}(v_{1,\ell},\ldots,v_{n,\ell})\Psi^{\langle\sigma\rangle}(v_{1,\ell},\ldots,v_{n,\ell}) = 0, \text{for } 0 \le \ell \le L_\times^{\langle\sigma\rangle} - 1,$$
$$L_\times^{\langle\sigma\rangle} = |D_{g,E;n_1} \times D_{f,E;n_1}^{\langle\sigma\rangle}|,\ \mathrm{supp}(\Phi^{\langle\sigma\rangle}) \subseteq D_{f,E;n_1}^{\langle\sigma\rangle}, \mathrm{supp}(\Psi^{\langle\sigma\rangle}) \subseteq D_{g,E;n_1} \quad (13)$$

satisfy

$$\Phi^{\langle\sigma\rangle} g - f^{\langle\sigma\rangle}\Psi^{\langle\sigma\rangle} = 0, \quad \mathrm{supp}(\Phi^{\langle\sigma\rangle}) \subseteq D_{f,E;n_1}^{\langle\sigma\rangle}, \mathrm{supp}(\Psi^{\langle\sigma\rangle}) \subseteq D_{g,E;n_1}. \quad (14)$$

That linear system (13) has $M^{\langle\sigma\rangle} = |D_{g,E;n_1}| + |D_{f,E;n_1}^{\langle\sigma\rangle}|$ unknown coefficients of $(\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle})$, and since we assume $k = 0$, that is, there are no erroneous evaluation ($\Lambda = 1$), in the supports (4) we have at least $E+1$ linearly independent solutions: $(f^{\langle\sigma\rangle} x_{n_1}^\nu, g\, x_{n_1}^\nu)$ with $0 \le \nu \le E$. We now can select $r_v^{\langle\sigma\rangle} \le M^{\langle\sigma\rangle} - E - 1$ linearly independent equations and preserve the solution. The crucial argument is that each of those rows is formed from a vector of new variables $(v_{1,\ell},\ldots,v_{n,\ell})$, so by variable substitution we conclude that the first $r_v^{\langle\sigma\rangle}$ rows in (13) already have maximal rank and yield (14), that over any field extension of $\mathsf{K}(v_{1,0},\ldots,v_{n,r_v^{\langle\sigma\rangle}})$. Here we use $k = 0$: for symbolic evaluations there are no (true) poles, so all equations have the same form.

The case $s \ge 1$ and $k = 0$ follows by using the first $\max_\sigma r_v^{\langle\sigma\rangle} \le (\max_\sigma M^{\langle\sigma\rangle}) - E - 1$ equations in (13) simultaneously for all σ. Note that for each individual σ the block of equations (13) produces solutions with a $\Psi^{\langle\sigma\rangle}$ component that satisfies (14), which are restricted to $\Psi^{\langle 1\rangle} = \cdots = \Psi^{\langle s\rangle}$, and (13) remains valid for the restriction. □

We return to the proof of Theorem 1.1.

We seek conditions on the $\xi_{\mu,\ell} \in S \subseteq \mathsf{K}$ so that the interpolants of (5) do not form a proper superspace to (8)–(11) of Lemma 2.1. There are at least $L - E = L_+ = |D_{g,E;n_1}| + (\max_{1\le\sigma\le s} |D_{f,E;n_1}^{\langle\sigma\rangle}|) - E - 1$ equations in (5) without erroneous $\beta_\ell^{\langle\sigma\rangle}$. Let $0 \le \ell_1 < \cdots < \ell_{L_+} \le L - 1$ be indices for good evaluations. Note that the ℓ_θ are not known on input.

By $r_v^{\langle\sigma\rangle} \le M^{\langle\sigma\rangle} - E - 1 \le L_+$ we have denoted for a given σ the rank of the symbolic (generic) interpolation system (13) in the proof of Lemma 2.3, which is the minimum number of symbolic equations necessary to obtain the Welch-Berlekamp property (14). For each σ we consider 4 solution spaces:

$$Y_\xi^{\langle\sigma\rangle} = \{(\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle}) \mid \Phi^{\langle\sigma\rangle}(\xi_{1,\ell_\theta},\ldots,\xi_{n,\ell_\theta}) - \beta_{\ell_\theta}^{\langle\sigma\rangle}\Psi^{\langle\sigma\rangle}(\xi_{1,\ell_\theta},\ldots,\xi_{n,\ell_\theta}) = 0,$$
$$\text{for all } 1 \le \theta \le r_v^{\langle\sigma\rangle}, \beta_{\ell_\theta}^{\langle\sigma\rangle} \ne \infty, \quad (15)$$
$$\Psi^{\langle\sigma\rangle}(\xi_{1,\ell_\theta},\ldots,\xi_{n,\ell_\theta}) = 0, \ \text{for all } 1 \le \theta \le r_v^{\langle\sigma\rangle}, \beta_{\ell_\theta}^{\langle\sigma\rangle} = \infty, \quad (16)$$
$$\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle} \in \mathsf{K}[x_1,\ldots,x_n],$$
$$\mathrm{supp}(\Phi^{\langle\sigma\rangle}) \subseteq D_{f,E;n_1}^{\langle\sigma\rangle},\ \mathrm{supp}(\Psi^{\langle\sigma\rangle}) \subseteq D_{g,E;n_1}\},$$

$$Z_{\mathsf{K}}^{\langle\sigma\rangle} = \{(\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle}) \mid \Phi^{\langle\sigma\rangle} g - f^{\langle\sigma\rangle}\Psi^{\langle\sigma\rangle} = 0, \quad \Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle} \in \mathsf{K}[x_1, \ldots, x_n], \\ \operatorname{supp}(\Phi^{\langle\sigma\rangle}) \subseteq D_{f,E;n_1}^{\langle\sigma\rangle}, \ \operatorname{supp}(\Psi^{\langle\sigma\rangle}) \subseteq D_{g,E;n_1}\}, \tag{17}$$

$$Y_v^{\langle\sigma\rangle} = \{(\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle}) \mid g(v_{1,\ell_\theta}, \ldots, v_{n,\ell_\theta})\Phi^{\langle\sigma\rangle}(v_{1,\ell_\theta}, \ldots, v_{n,\ell_\theta}) - \\ f^{\langle\sigma\rangle}(v_{1,\ell_\theta}, \ldots, v_{n,\ell_\theta})\Psi(v_{1,\ell_\theta}, \ldots, v_{n,\ell_\theta}) = 0 \text{ for all } 1 \le \theta \le r_v^{\langle\sigma\rangle}, \\ \Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle} \in \mathsf{K}(v_{1,\ell_1}, \ldots, v_{n,\ell_\rho})[x_1, \ldots, x_n], \rho = r_v^{\langle\sigma\rangle}, \\ \operatorname{supp}_{x_1,\ldots,x_n}(\Phi^{\langle\sigma\rangle}) \subseteq D_{f,E;n_1}^{\langle\sigma\rangle}, \ \operatorname{supp}_{x_1,\ldots,x_n}(\Psi^{\langle\sigma\rangle}) \subseteq D_{g,E;n_1}\}, \tag{18}$$

$$Z_{\mathsf{K}(v)}^{\langle\sigma\rangle} = \{(\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle}) \mid \Phi^{\langle\sigma\rangle} g - f^{\langle\sigma\rangle}\Psi^{\langle\sigma\rangle} = 0, \\ \Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle} \in \mathsf{K}(v_{1,\ell_1}, \ldots, v_{n,\ell_\rho})[x_1, \ldots, x_n], \rho = r_v^{\langle\sigma\rangle}, \\ \operatorname{supp}(\Phi^{\langle\sigma\rangle}) \subseteq D_{f,E;n_1}^{\langle\sigma\rangle}, \ \operatorname{supp}(\Psi^{\langle\sigma\rangle}) \subseteq D_{g,E;n_1}\}. \tag{19}$$

Note that (18) has the equations of (13) multiplied by the polynomial denominator of $\beta_{\ell_\theta}^{\langle\sigma\rangle}(v_{1,\ell_\theta}, \ldots, v_{n,\ell_\theta})$, namely $g(v_{1,\ell_\theta}, \ldots, v_{n,\ell_\theta})$. The coefficient matrix of (18), which has $r_v^{\langle\sigma\rangle}$ rows, has full rank $r_v^{\langle\sigma\rangle}$ and there is a non-singular $r_v^{\langle\sigma\rangle} \times r_v^{\langle\sigma\rangle}$ submatrix, denoted by $A_v^{\langle\sigma\rangle}(v_{1,\ell_1}, \ldots, v_{n,\ell_\rho})$, whose determinant is not zero,

$$\Delta_v^{\langle\sigma\rangle}(v_{1,\ell_1}, \ldots, v_{n,\ell_\rho}) = \det(A_v^{\langle\sigma\rangle}(v_{1,\ell_1}, \ldots, v_{n,\ell_\rho})) \neq 0, \quad \rho = r_v^{\langle\sigma\rangle}. \tag{20}$$

We shall assume that for ξ_{μ,ℓ_θ}, $1 \le \mu \le n$, $1 \le \theta \le r_v^{\langle\sigma\rangle}$ we have

$$\Delta_v^{\langle\sigma\rangle}(\xi_{1,\ell_1}, \ldots, \xi_{n,\ell_\rho}) \neq 0, \quad \rho = r_v^{\langle\sigma\rangle}. \tag{21}$$

We now apply Lemma 2.1 for $s = 1$ to (17) and (15, 16), which have $L = r_v^{\langle\sigma\rangle}$ and $k = 0$, the latter of which renders (9) and (10) vacuous. Because by (21) no row in $A_v^{\langle\sigma\rangle}(\xi_{1,\ell_1}, \ldots, \xi_{n,\ell_\rho})$ can be a zero row, we have

$$\forall\, \theta, 1 \le \theta \le r_v^{\langle\sigma\rangle} \colon f^{\langle\sigma\rangle}(\xi_{1,\ell_\theta}, \ldots, \xi_{n,\ell_\theta}) \neq 0 \quad \text{or} \quad g(\xi_{1,\ell_\theta}, \ldots, \xi_{n,\ell_\theta}) \neq 0. \tag{22}$$

So (11) is also vacuous. Thus Lemma 2.1 applies, and we obtain $Y_\xi^{\langle\sigma\rangle} \supseteq Z_{\mathsf{K}}^{\langle\sigma\rangle}$.

By (21) the coefficient matrix of (15,16) also has maximal rank $r_v^{\langle\sigma\rangle}$: a special case are the rows corresponding to the equations (16) for true poles $g(\xi_{1,\ell_\theta}, \ldots, \xi_{n,\ell_\theta}) = 0$, which in $A_v^{\langle\sigma\rangle}(\xi_{1,\ell_1}, \ldots, \xi_{n,\ell_\rho})$ are multiplied by $f^{\langle\sigma\rangle}(\xi_{1,\ell_\theta}, \ldots, \xi_{n,\ell_\theta})$, which by (22) is a non-zero scalar in K. Therefore, the vector space dimension of $Y_\xi^{\langle\sigma\rangle}$ is $M^{\langle\sigma\rangle} - r_v^{\langle\sigma\rangle}$, where $M^{\langle\sigma\rangle} = |D_{g,E;n_1}| + |D_{f,E;n_1}^{\langle\sigma\rangle}|$ is the number of unknown coefficients in (15,16). The vector space dimension of $Z_{\mathsf{K}}^{\langle\sigma\rangle}$ is equal to the dimension of $Z_{\mathsf{K}(v)}^{\langle\sigma\rangle}$, because the entries of the coefficient matrix of (17) and (19) depend only on $f^{\langle\sigma\rangle}$ and g. The vector space dimension of $Y_v^{\langle\sigma\rangle}$ is $M^{\langle\sigma\rangle} - r_v^{\langle\sigma\rangle}$ by definition of $r_v^{\langle\sigma\rangle}$, and $Y_v^{\langle\sigma\rangle} = Z_{\mathsf{K}(v)}^{\langle\sigma\rangle}$ by the arguments in the proof of Lemma 2.3, so $Z_{\mathsf{K}}^{\langle\sigma\rangle}$ also has dimension $M^{\langle\sigma\rangle} - r_v^{\langle\sigma\rangle}$. Since $Y_\xi^{\langle\sigma\rangle} \supseteq Z_{\mathsf{K}}^{\langle\sigma\rangle}$ and the two vector spaces have the same dimension, we finally get $Y_\xi^{\langle\sigma\rangle} = Z_{\mathsf{K}}^{\langle\sigma\rangle}$.

Next, we restrict the solutions $(\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle}) \in Y_\xi^{\langle\sigma\rangle}$ by all remaining equational constraints for true evaluations in (5):

$$\Phi^{\langle\sigma\rangle}(\xi_{1,\ell},\ldots,\xi_{n,\ell}) - \beta_\ell^{\langle\sigma\rangle}\Psi^{\langle\sigma\rangle}(\xi_{1,\ell},\ldots,\xi_{n,\ell}) = 0, \quad \text{for all } 0 \le \ell \le L-1 \text{ with}$$
$$\ell \ne \ell_\theta \text{ (for all } 1 \le \theta \le r_v^{\langle\sigma\rangle}), \ell \ne \lambda_\kappa \text{ (for all } 1 \le \kappa \le k), \beta_\ell^{\langle\sigma\rangle} \ne \infty, \tag{23}$$
$$\Psi^{\langle\sigma\rangle}(\xi_{1,\ell},\ldots,\xi_{n,\ell}) = 0, \quad \text{for all } 1 \le \ell \le L \text{ with}$$
$$\ell \ne \ell_\theta \text{ (for all } 1 \le \theta \le r_v^{\langle\sigma\rangle}), \ell \ne \lambda_\kappa \text{ (for all } 1 \le \kappa \le k), \beta_\ell^{\langle\sigma\rangle} = \infty, \tag{24}$$

Because $Y_\xi^{\langle\sigma\rangle} = Z_{\mathsf{K}}^{\langle\sigma\rangle}$, all $(\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle})$ satisfy (17) and therefore (23) and (24), the latter provided $f^{\langle\sigma\rangle}(\xi_{1,\ell},\ldots,\xi_{n,\ell}) \ne 0$. If $f^{\langle\sigma\rangle}(\xi_{1,\ell},\ldots,\xi_{n,\ell}) = g(\xi_{1,\ell},\ldots,\xi_{n,\ell}) = 0$, the constraints (24) restrict the solution space, but (17) remains valid for the solutions.

We conclude as at the end of the proof of Lemma 2.3, assuming that the condition (21) is simultaneously satisfied for all σ. The system (5), excluding all erroneous $\ell \in \{\lambda_1,\ldots,\lambda_k\}$, solves (15) simultaneously for all σ and a common denominator $\Psi^{\langle 1\rangle} = \cdots = \Psi^{\langle s\rangle}$, which restricts the solution pairs $(\Phi^{\langle\sigma\rangle}, \Psi^{\langle\sigma\rangle})$ to a further subspace of $Z_{\mathsf{K}}^{\langle\sigma\rangle}$, and the Welch-Berlekamp property (8) remains valid.

We finally include all remaining erroneous equations in (6) at $\ell \in \{\lambda_1,\ldots,\lambda_k\}$. From the determinantal conditions (21) for all σ, as we have just shown, all solutions $(\Phi^{\langle 1\rangle},\ldots,\Phi^{\langle s\rangle},\Psi)$ of the (good) equations in (6) at all ℓ with $0 \le \ell \le L-1$ and $\ell \notin \{\lambda_1,\ldots,\lambda_k\}$ satisfy the Welch-Berlekamp property (8), which is (17) for all σ. Hence we must have

$$\forall\sigma\colon (\Phi^{\langle\sigma\rangle} g)(\xi_{1,\lambda_\kappa},\ldots) = (f^{\langle\sigma\rangle}\Psi)(\xi_{1,\lambda_\kappa},\ldots). \tag{25}$$

The equations with erroneous $\beta_{\lambda_\kappa}^{\langle\sigma\rangle} \ne \infty$ have

$$\forall\sigma\colon \Phi^{\langle\sigma\rangle}(\xi_{1,\lambda_\kappa},\ldots) = \beta_{\lambda_\kappa}^{\langle\sigma\rangle}\Psi(\xi_{1,\lambda_\kappa},\ldots), \quad \exists\sigma_1\colon f^{\langle\sigma_1\rangle}(\xi_{1,\lambda_\kappa},\ldots) \ne \beta_{\lambda_\kappa}^{\langle\sigma_1\rangle} g(\xi_{1,\lambda_\kappa},\ldots). \tag{26}$$

Thus,

$$(f^{\langle\sigma_1\rangle}\Psi)(\xi_{1,\lambda_\kappa},\ldots) = (\Phi^{\langle\sigma_1\rangle} g)(\xi_{1,\lambda_\kappa},\ldots) = \beta_{\lambda_\kappa}^{\langle\sigma_1\rangle}(\Psi g)(\xi_{1,\lambda_\kappa},\ldots),$$

which by (26) forces

$$\forall\sigma\colon \Psi(\xi_{1,\lambda_\kappa},\ldots) = 0 = \Phi^{\langle\sigma\rangle}(\xi_{1,\lambda_\kappa},\ldots). \tag{27}$$

For the equations with erroneous $\beta_{\lambda_\kappa}^{\langle\sigma\rangle} = \infty$ we have the equations $\Psi(\xi_{1,\lambda_\kappa},\ldots) = 0$, but $g(\xi_{1,\lambda_\kappa},\ldots) \ne 0$, which with (25) completes (27). We conclude that all erroneous equations restrict the interpolation solution space by (27), and the Welch-Berlekamp property (8) stays preserved. Again, there is at least one non-zero solution $(f^{\langle 1\rangle}\Lambda,\ldots,f^{\langle s\rangle}\Lambda, g\Lambda)$.

The probabilistic analysis for condition (21) uses the DeMillo-Lipton-Schwartz-Zippel Lemma. Let $H = \prod_{\sigma=1}^s \Delta_v^{\langle\sigma\rangle}$. Then $H(\xi_{1,\ell_1},\ldots,\xi_{n,\ell_\rho}) \ne 0$, $\rho = \max_\sigma\{r_v^{\langle\sigma\rangle}\}$, yields (21) for all σ, that for uniformly sampled random $\xi_{\mu,\ell_\theta} \in S \subseteq \mathsf{K}$ with a probability $\ge 1 - \deg(H)/|S|$, where $|S|$ is the number of elements in the finite set S. Note that the black box can produce erroneous equations adaptively to the evaluation points $(\xi_{1,\ell},\ldots,\xi_{n,\ell})$, akin to the adaptive cipher text attack in public key crypto systems. But our evaluations are random for each ℓ, so they are random for those (unknown) equations at ℓ_θ in (21).

The proof of Theorem 1.1 concludes by bounding $\deg(H) = \sum_{\sigma=1}^s \deg(\Delta_v^{\langle\sigma\rangle})$. Each entry in the coefficient matrix $A_v^{\langle\sigma\rangle}$ in (20) is either a term in $D_{f,E;n_1}^{\langle\sigma\rangle}$ at the variables $v_{1,\ell_\theta},\ldots,v_{n,\ell_\theta}$ $\times$

$g(v_{1,\ell_\theta},\ldots,v_{n,\ell_\theta})$ or a term in $D_{g,E;n_1}$ at the variables $v_{1,\ell_\theta},\ldots,v_{n,\ell_\theta} \times (-f^{\langle\sigma\rangle}(v_{1,\ell_\theta},\ldots,v_{n,\ell_\theta}))$. We have $\deg(f^{\langle\sigma\rangle}) \le \max\{\deg(\tau_f) \mid \tau_f \in D_f^{\langle\sigma\rangle}\} = \max\{\deg(\tau_f) \mid \tau_f \in D_{f,E;n_1}\} - E$ and $\deg(g) \le \max\{\deg(\tau_g) \mid \tau_g \in D_g\} = \max\{\deg(\tau_g) \mid \tau_g \in D_{g,E;n_1}\} - E$. Therefore, in (20) we have

$$\begin{aligned}\deg(\Delta_v^{\langle\sigma\rangle}) &\le r_v^{\langle\sigma\rangle} \times \max\{\deg(f^{\langle\sigma\rangle}) + \max\{\deg(\tau_g) \mid \tau_g \in D_{g,E;n_1}\},\\ &\qquad\qquad \deg(g) + \max\{\deg(\tau_f) \mid \tau_f \in D_{f,E;n_1}^{\langle\sigma\rangle}\} \qquad \}\\ &\le (M^{\langle\sigma\rangle} - E - 1) \times (\max\{\deg(\tau_f) \mid \tau_f \in D_f^{\langle\sigma\rangle}\} + \max\{\deg(\tau_g) \mid \tau_g \in D_g\} + E).\end{aligned}$$

Thus ends the proof of Theorem 1.1. □

Remark 2.2. The condition $g \neq 0$ in (1) is not essential, which is useful when we inadvertently have projected the denominator to 0 during a Zipple-sytle iteration (Kaltofen and Yang 2014, Section 3). If $g = 0$, all non-faulty evaluation vectors are by definition $[\infty,\ldots,\infty]$. If for all σ: $D_f^{\langle\sigma\rangle} = \varnothing \Rightarrow f^{\langle\sigma\rangle} = 0$, then the Welch-Berlekamp property (6) is satisfied for any solution (Φ,Ψ). Otherwise, if for one σ_1 we have $D_f^{\langle\sigma_1\rangle} \neq \varnothing$, we must have $\Psi = 0$ for all solutions of (5). We obtain $\Psi = 0$ for the symbolic evaluations in Lemma 2.2: from $g = 0$ we have $\beta_\ell^{\langle\sigma\rangle} = \infty$ for all ℓ and the (shortened) (13) is $\Psi^{\langle\sigma_1\rangle}(v_{1,\ell},\ldots,v_{n,\ell}) = 0$, for $0 \le \ell \le L_+ - 1$, where $L_+ \ge |D_{f,E;n_1}^{\langle\sigma_1\rangle}| + |D_{g,E;n_1}| - E - 1 \ge |D_{g,E;n_1}|$, the latter because $D_{f,E;n_1}^{\langle\sigma_1\rangle}$ contains at least $E+1$ terms. Therefore $\Psi^{\langle\sigma_1\rangle} = 0 = \Psi$; Theorem 1.1 follows as in the rest of the proof. □

Remark 2.3. The system of linear equations (6) is not entirely square: there can be as many as $s \times L$ equations in at most $(s \times \max_{1\le\sigma\le s} |D_{f,E;n_1}^{\langle\sigma\rangle}|) + |D_{g,E;n_1}|$ unknown coefficients of the $\Phi^{\langle\sigma\rangle}$ and Ψ. With the bound (7) of L those are $(s-1)|D_{g,E;n_1}|$ more equations than unknows. If $f^{\langle 1\rangle} = \cdots = f^{\langle s\rangle}$ and $\beta_\ell^{\langle 1\rangle} = \cdots = \beta_\ell^{\langle s\rangle}$ for all ℓ, the number of evaluations L constitutes the case $s = 1$, which is a square system, and therefore cannot be reduced. However, reduction can be possible for dense univariate rational function recovery (see Remark 2.1) when the vector of rational functions is the solution of a linear system (Cabay 1971; Olesh and Storjohann 2007). In (Pernet 2014, Section 2.4) such reductions are cited for collaborative decoding Reed-Solomon codes (Schmidt et al. 2006) under genericity assumption of the vector of fractions, and for solutions to linear systems the precise condition is known (Kaltofen et al. 2015). For sparse multivariate rational function recovery and decoding, we do not know what genericity would yield reduction to $L = (\max_{1\le\sigma\le s} |D_{f,E;n_1}^{\langle\sigma\rangle}|) + \lceil |D_{g,E;n_1}|/s \rceil$. □

Remark 2.4. Theorem 1.1 and Lemma 1.2 state the probability of obtaining maximal rank for exact arithmetic. With floating point arithmetic, a lower bound of the expected condition number or first non-zero singular value is necessitated in the probabilistic analysis. The numeric counterpart of the DeMillo-Lipton-Schwartz-Zippel Lemma, namely Lemma 3.1 in (Kaltofen et al. 2007), can be applied to that task. □

Acknowledgements: We thank Clément Pernet for comments inducing Remark 2.3 and Terence Tao for comments inducing Lemma 1.2.

Note added on March 16, 2016: added DeMillo-Lipton to the "DeMillo-Lipton-Schwartz-Zippel Lemma."

References

Cabay, Stanley. Exact solution of linear equations. In *Proceedings of the Second ACM Symposium on Symbolic and Algebraic Manipulation*, SYMSAC '71, pages 392–398, New York, NY, USA, 1971. ACM. URL http://doi.acm.org/10.1145/800204.806310.

Kaltofen, Erich and Pernet, Clément. Cauchy interpolation with errors in the values. Manuscript, 13 pages, December 2013.

Kaltofen, Erich L., Pernet, Clément, Storjohann, Arne, and Waddell, Cleveland A. Linear system solving with parametric entries by error correction and Cabay termination. Poster at the Internat. Symp. Symbolic Algebraic Comput. 2015, 2015. Paper in preparation.

Kaltofen, Erich and Yang, Zhengfeng. Sparse multivariate function recovery from values with noise and outlier errors. In Kauers, Manuel, editor, *ISSAC 2013 Proc. 38th Internat. Symp. Symbolic Algebraic Comput.*, pages 219–226, New York, N. Y., 2013. Association for Computing Machinery. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/13/KaYa13.pdf.

Kaltofen, Erich L. and Yang, Zhengfeng. Sparse multivariate function recovery with a high error rate in evaluations. In Nabeshima, Katsusuke, editor, *ISSAC 2014 Proc. 39th Internat. Symp. Symbolic Algebraic Comput.*, pages 280–287, New York, N. Y., 2014. Association for Computing Machinery. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/14/KaYa14.pdf.

Kaltofen, Erich, Yang, Zhengfeng, and Zhi, Lihong. On probabilistic analysis of randomization in hybrid symbolic-numeric algorithms. In Verschelde, Jan and Watt, Stephen M., editors, *SNC'07 Proc. 2007 Internat. Workshop on Symbolic-Numeric Comput.*, pages 11–17, New York, N. Y., 2007. ACM Press. ISBN 978-1-59593-744-5. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/07/KYZ07.pdf.

Olesh, Zach and Storjohann, Arne. The vector rational function reconstruction problems. In *Proc. Waterloo Workshop on Computer Algebra: devoted to the 60th birthday of Sergei Abramov (WWCA)*, pages 137–149, 2007.

Olshevsky, V. and Shokrollahi, M. Amin. A displacement approach to decoding algebraic codes. In *Algorithms for Structured Matrices: Theory and Applications*, pages 265–292. American Mathematical Society, Providence, Rhode Island, USA, 2003. Contemporary Math., vol. 323. URL: http://www.math.uconn.edu/~olshevsky/papers/shokrollahi_f.pdf.

Pernet, Clément. *High Performance and Reliable Algebraic Computing.* Mémoire d'habilitation à diriger des recherches, Université Joseph Fourier (Grenoble 1), November 2014.

Schmidt, Georg, Sidorenko, Vladimir, and Bossert, Martin. Collaborative decoding of interleaved reed-solomon codes and concatenated code designs. *CoRR*, abs/cs/0610074, 2006. URL http://arxiv.org/abs/cs/0610074.

Tao, Terence. An uncertainty principle for cyclic groups of prime order. *Math. Research Letters*, 12:121–127, 2005. URL:http://arxiv.org/pdf/math/0308286.pdf.

Zippel, R. E. *Probabilistic algorithms for sparse polynomials.* PhD thesis, Massachusetts Inst. of Technology, Cambridge, USA, September 1979.

3. Appendix

Notation (in alphabetic order):	
β	the possibly erroneous values returned by the black box for f/g
γ	the correct evaluations for f/g
$d_{j,\mu}$	the degree of variable x_μ in the j-th term in f
$\bar{d}_f$	$\geq \deg(f)$, bounds that are input
$\bar{d}_g$	$\geq \deg(g)$, bounds that are input
D	sets of terms (non-zero monomials)
Δ	a matrix determinant
$e_{m,\mu}$	the degree of variable x_μ in the m-th term in g
E	$\geq k$, an upper bound on the number of errors that is input to the algorithm
$f, f^{\langle\sigma\rangle}$	the numerator polynomial, polynomials with $1 \leq \sigma \leq s$
$\Phi, \Phi^{\langle\sigma\rangle}$	the numerator(s) of the computed interpolant
g	the (common) denominator polynomial
k	the actual number of errors, to be determined by the algorithms
K	an arbitrary field with exact arithmetic
L	the length of the list of a batch of evaluations
λ_κ	$1 \leq \kappa \leq k$, are the positions of the erroneous evaluations in the list of evaluations
Λ	the error locator polynomial
M	the number of unknowns in our linear systems
μ	a subscript that corresponds to the μ-th variable x_μ, $1 \leq \mu \leq n$
n	is the number of variables
n_1	$\Lambda(x_{n_1})$ is the univariate error locator polynomial with $1 \leq n_1 \leq n$
Ψ	the common denominator of the computed interpolant
q	$1/q$ is the error rate;
r	the rank of a matrix
s	the number of fractions in vector recovery
$\mathrm{supp}(f)$	support of f: the set of terms $\{x_1^{d_{1,j}} \cdots x_n^{d_{n,j}} \mid j = 1, \ldots, t\}$ in $f = \sum_{j=1}^{t} a_j x_1^{d_{1,j}} \cdots x_n^{d_{n,j}}$ occurring with non-zero coefficients $\forall j \colon a_j \neq 0$.
σ	a component in a vector of fractions, $1 \leq \sigma \leq s$
τ	a placeholder symbol for any term $\in D$: τ_f for terms in the term-supersets of f, τ_g for terms in the term-supersets of g.
t	denotes the actual number of terms
T	$\geq t$, an upper bound that is input
x_i	the variables of f/g
ξ	values for the variables from a field $\in \mathsf{K}$
Y	the vector space of interpolants
Z	the vector space of equal fractions