Correspondence

A Comment on the Weiss–Weinstein Bound for Constrained Parameter Sets

Zvika Ben-Haim and Yonina C. Eldar, *Senior Member, IEEE*

***Abstract*—The Weiss–Weinstein bound (WWB) provides a lower limit on the mean-squared error (MSE) achievable by an estimator of an unknown random parameter. In this correspondence, it is shown that some previously proposed simplified versions of the bound do not always hold for constrained parameters, i.e., parameters whose distribution has finite support. These simplifications can produce results which are no longer lower bounds on the MSE. Sufficient conditions are provided for the reductions to be valid.**

***Index Terms*—Bayesian estimation, constrained estimation, performance bounds, Weiss–Weinstein bound.**

I. Introduction

We consider the problem of estimating a random vector $\boldsymbol{\theta}$ from observations $\boldsymbol{x}$, where the quality of an estimator $\boldsymbol{g}(\boldsymbol{x})$ is measured by its mean-squared error (MSE) $E\left\{\|\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{\theta}\|^2\right\}$. It is well known that the posterior mean $E\{\boldsymbol{\theta}|\boldsymbol{x}\}$ is the technique minimizing the MSE. However, in many cases, determining the posterior mean is computationally prohibitive, and various approaches have been developed as alternatives. It is therefore of interest to determine the degradation in accuracy resulting from the use of suboptimal methods. Unfortunately, computation of the optimal MSE is itself infeasible in many cases. This has led to a large body of work seeking to find simple lower bounds for the minimum MSE in a given estimation problem [1]–[4].

In a landmark paper, Weiss and Weinstein developed a general technique for deriving lower bounds on the minimum achievable MSE [3]. A noteworthy feature of their method is that it requires almost no regularity assumptions on the problem setting. The bound was further developed in [4], and has been used in a variety of practical applications [5]–[7].

A common setting in which the Weiss–Weinstein bound (WWB) is often applied concerns constrained parameter sets, i.e., situations in which the parameter $\boldsymbol{\theta}$ occurs with probability 1 in a subset Θ of $\mathbb{R}^m$. For example, in time-delay estimation, the delay is sometimes assumed to be uniformly distributed in a given interval [3]. While the WWB continues to hold in the constrained setting, some of the simplifications presented in [3] and [4] are not valid in this case.

In this correspondence, we point out the versions of the WWB which do not necessarily hold for constrained parameter sets. We discuss the regularity conditions under which these simplified versions are valid, and conclude with an example in which the simplified version of the bound yields incorrect (and unreasonable) results.

Manuscript received November 6, 2007; revised June 29, 2008. Current version published September 17, 2008.

The authors are with the Department of Electrical Engineering, Technion–Israel Institute of Technology, Technion City, Haifa 32000, Israel (e-mail: zvikabh@tx.technion.ac.il; yonina@ee.technion.ac.il).

Communicated by L. Tong, Associate Editor for Detection and Estimation.
Digital Object Identifier 10.1109/TIT.2008.928970

II. Background and Summary

We begin by recalling some definitions and results relating to the WWB [3], [4]. Let $\boldsymbol{\theta}$ and $\boldsymbol{x}$ be finite-variance random vectors whose joint probability density function (pdf) is $f(\boldsymbol{x},\boldsymbol{\theta})$. Suppose that $f(\boldsymbol{x},\boldsymbol{\theta})$ is nonzero only for values of $\boldsymbol{\theta}$ in a subset Θ of $\mathbb{R}^m$, and let Θ' be the set of values of $\boldsymbol{\theta}$ for which $f(\boldsymbol{x},\boldsymbol{\theta})$ is positive almost everywhere (a.e.) in $\boldsymbol{x}$. We are interested in estimating $\boldsymbol{\theta}$ using a function $\boldsymbol{g}(\boldsymbol{x})$ of the measurements. The error covariance matrix is defined as

$$\boldsymbol{R} \triangleq E\left\{(\boldsymbol{\theta}-\boldsymbol{g}(\boldsymbol{x}))(\boldsymbol{\theta}-\boldsymbol{g}(\boldsymbol{x}))^T\right\}. \tag{1}$$

The goal is to find a lower bound on $\boldsymbol{R}$, i.e., a matrix $\boldsymbol{B}$ such that $\boldsymbol{R} \geq \boldsymbol{B}$, where the matrix inequality means that $\boldsymbol{R}-\boldsymbol{B}$ is positive semidefinite.

Denote the likelihood ratio by

$$L(\boldsymbol{x};\boldsymbol{\theta}_1,\boldsymbol{\theta}_2) \triangleq \frac{f(\boldsymbol{x},\boldsymbol{\theta}_1)}{f(\boldsymbol{x},\boldsymbol{\theta}_2)} \tag{2}$$

where the function is defined only for values of $\boldsymbol{x}$ and $\boldsymbol{\theta}_2$ such that $f(\boldsymbol{x},\boldsymbol{\theta}_2) \neq 0$. Let

$$\mu(s,\boldsymbol{h}) \triangleq \ln E\left\{L^s(\boldsymbol{x};\boldsymbol{\theta}+\boldsymbol{h},\boldsymbol{\theta})\right\} \tag{3}$$

and note that the expectation is calculated only over points $(\boldsymbol{x},\boldsymbol{\theta})$ such that $f(\boldsymbol{x},\boldsymbol{\theta})>0$, i.e., those points for which $L(\boldsymbol{x};\boldsymbol{\theta}+\boldsymbol{h},\boldsymbol{\theta})$ is defined.

In the case of a scalar parameter $\theta \in \Theta \subseteq \mathbb{R}$, the error covariance (1) equals the MSE. The WWB in this setting is given by

$$\text{MSE} \geq \frac{h^2 E^2\{L^s(\boldsymbol{x};\theta+h,\theta)\}}{E\left\{\left[L^s(\boldsymbol{x};\theta+h,\theta)-L^{1-s}(\boldsymbol{x};\theta-h,\theta)\right]^2\right\}} \tag{4}$$

for any h and s such that

$$0 < E\left\{\left[L^s(\boldsymbol{x};\theta+h,\theta)-L^{1-s}(\boldsymbol{x};\theta-h,\theta)\right]^2\right\} < \infty. \tag{5}$$

If Θ' is a connected subset of $\mathbb{R}$, then (4) can also be written as [3], [4], [7]

$$\text{MSE} \geq \frac{h^2 e^{2\mu(s,h)}}{e^{\mu(2s,h)}+e^{\mu(2-2s,-h)}-2e^{\mu(s,2h)}}. \tag{6}$$

In these cases, calculation of $\mu(s,h)$ is sufficient for evaluation of the bound. However, the equivalence between the original bound (4) and the simplified version (6) does not necessarily hold if Θ' is a disjoint subset of $\mathbb{R}$; specifically, the cross-term $2e^{\mu(s,2h)}$ does not always equal $2E\left\{L^s(\boldsymbol{x};\theta+h,\theta)L^{1-s}(\boldsymbol{x};\theta-h,\theta)\right\}$. As we will see, when Θ' is disjoint, (6) can be larger than the minimum estimation MSE, and may even be infinite. Thus, care must be used when applying the bound to disjoint parameter sets Θ'.

In the case of a vector parameter $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^m$, the WWB on the error covariance is given by

$$\boldsymbol{R} \geq \boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{H}^T. \tag{7}$$

Here, $\boldsymbol{H}=[\boldsymbol{h}_1,\boldsymbol{h}_2,\ldots,\boldsymbol{h}_m]$ is a matrix consisting of "test vectors" $\boldsymbol{h}_i \in \mathbb{R}^m$, and $\boldsymbol{G}$ is the $m \times m$ matrix whose elements are given by

$$G_{ij} = \frac{E\left\{r(\boldsymbol{x},\boldsymbol{\theta};\boldsymbol{h}_i,s_i)r(\boldsymbol{x},\boldsymbol{\theta};\boldsymbol{h}_j,s_j)\right\}}{E\left\{L^{s_i}(\boldsymbol{x};\boldsymbol{\theta}+\boldsymbol{h}_i)\right\}E\left\{L^{s_j}(\boldsymbol{x};\boldsymbol{\theta}+\boldsymbol{h}_j)\right\}} \tag{8}$$

where

$$r(\boldsymbol{x},\boldsymbol{\theta};\boldsymbol{h}_i,s_i) \triangleq L^{s_i}(\boldsymbol{x};\boldsymbol{\theta}+\boldsymbol{h}_i,\boldsymbol{\theta}) - L^{1-s_i}(\boldsymbol{x};\boldsymbol{\theta}-\boldsymbol{h}_i,\boldsymbol{\theta}). \quad (9)$$

The bound holds for any $\{\boldsymbol{h}_i,s_i\}_{i=1}^m$ such that $\boldsymbol{G}$ is well defined and invertible.

Weinstein and Weiss [4] suggest that attention be restricted to the case $s_1 = \cdots = s_m = 1/2$. In this case, and under the additional assumption that $\Theta' = \mathbb{R}^m$, the bound (8) simplifies to[1]

$$G_{ij} = 2\frac{e^{\mu(1/2,\boldsymbol{h}_i-\boldsymbol{h}_j)} - e^{\mu(1/2,\boldsymbol{h}_i+\boldsymbol{h}_j)}}{e^{\mu(1/2,\boldsymbol{h}_i)}e^{\mu(1/2,\boldsymbol{h}_j)}}. \quad (10)$$

While this simplification is valid if $f(\boldsymbol{x},\boldsymbol{\theta})$ is positive a.e. (i.e., if $\Theta' = \mathbb{R}^m$), it does not necessarily hold in other cases. Furthermore, when $\boldsymbol{\theta}$ is a scalar, (10) does not necessarily reduce to (6).

III. Scalar Case

To demonstrate equivalence between (4) and (6), it is required to show that

$$E\left\{L^s(\boldsymbol{x};\theta+h,\theta)L^{1-s}(\boldsymbol{x};\theta-h,\theta)\right\} \stackrel{?}{=} e^{\mu(s,2h)}. \quad (11)$$

When Θ' is a connected set, this statement can be verified as follows:

$$\begin{aligned} &E\left\{L^s(\boldsymbol{x};\theta+h,\theta)L^{1-s}(\boldsymbol{x};\theta-h,\theta)\right\} \\ &= \int \frac{f^s(\boldsymbol{x},\theta+h)}{f^s(\boldsymbol{x},\theta)}\frac{f^{1-s}(\boldsymbol{x},\theta-h)}{f^{1-s}(\boldsymbol{x},\theta)} f(\boldsymbol{x},\theta)d\boldsymbol{x}d\theta \quad (12)\\ &= \int f^s(\boldsymbol{x},\theta+h)f^{1-s}(\boldsymbol{x},\theta-h)d\boldsymbol{x}d\theta \quad (13)\\ &= \int \frac{f^s(\boldsymbol{x},\theta'+2h)}{f^s(\boldsymbol{x},\theta')}f(\boldsymbol{x},\theta')d\boldsymbol{x}d\theta' \quad (14)\\ &= e^{\mu(s,2h)} \quad (15)\end{aligned}$$

where we used the change of variables $\theta' = \theta - h$.

As noted previously, it is implicitly assumed in (11) that the expectation is calculated only over those values of $\boldsymbol{x}$ and θ for which $f(\boldsymbol{x},\theta) > 0$, otherwise $L(\boldsymbol{x};\theta+h,\theta)$ is undefined. Thus, the integral (12) is taken only over those values of $\boldsymbol{x}$ and θ for which $f(\boldsymbol{x},\theta) > 0$. However, after $f(\boldsymbol{x},\theta)$ is canceled out in (13), this fact is ignored. The integral (14) is taken over the range $f(\boldsymbol{x},\theta') > 0$, which corresponds to $f(\boldsymbol{x},\theta+h) > 0$. If there exist points $(\boldsymbol{x},\theta)$ for which $f(\boldsymbol{x},\theta) = 0$, $f(\boldsymbol{x},\theta+h) > 0$, and $f(\boldsymbol{x},\theta-h) > 0$, then (12) is not taken over those points, whereas (13) is positive at those points.

If Θ' is a connected set (i.e., a finite or infinite interval), then, for any $\boldsymbol{x}$ and θ such that $f(\boldsymbol{x},\theta-h) > 0$ and $f(\boldsymbol{x},\theta+h) > 0$, we also have $f(\boldsymbol{x},\theta) > 0$. In this case, the range of integration in (12) equals that of (13), so that the simplified version (6) is correct. This occurs, for instance, in the example given in [3], where Θ' is a closed interval.

However, if Θ' is disjoint, then (13) can be greater than (12), so that (6) can be larger than (4), and is not necessarily a lower bound on the MSE. Indeed, as we will see, in some cases (6) is higher than the minimum MSE; in other cases, (6) is infinite, as a result of a division by zero.

IV. Vector Case

When $\Theta' = \mathbb{R}^m$, the simplified equation (10) can be derived from the WWB (8) as follows. Substituting $s_1 = \cdots = s_m = 1/2$ in (8), the denominator equals that of (10). The numerator consists of a sum of four expressions of the type

$$E\left\{L^{1/2}(\boldsymbol{x};\boldsymbol{\theta}\pm\boldsymbol{h}_i,\boldsymbol{\theta})L^{1/2}(\boldsymbol{x};\boldsymbol{\theta}\pm\boldsymbol{h}_j,\boldsymbol{\theta})\right\}. \quad (16)$$

These can be simplified by writing

$$\begin{aligned} &E\left\{L^{1/2}(\boldsymbol{x};\boldsymbol{\theta}+\boldsymbol{h}_i,\boldsymbol{\theta})L^{1/2}(\boldsymbol{x};\boldsymbol{\theta}+\boldsymbol{h}_j,\boldsymbol{\theta})\right\} \\ &= \int\left(\frac{f(\boldsymbol{x},\boldsymbol{\theta}+\boldsymbol{h}_i)}{f(\boldsymbol{x},\boldsymbol{\theta})}\frac{f(\boldsymbol{x},\boldsymbol{\theta}+\boldsymbol{h}_j)}{f(\boldsymbol{x},\boldsymbol{\theta})}\right)^{1/2} f(\boldsymbol{x},\boldsymbol{\theta})d\boldsymbol{x}d\boldsymbol{\theta} \quad (17)\\ &= \int \left(f(\boldsymbol{x},\boldsymbol{\theta}+\boldsymbol{h}_i)f(\boldsymbol{x},\boldsymbol{\theta}+\boldsymbol{h}_j)\right)^{1/2}d\boldsymbol{x}d\boldsymbol{\theta} \quad (18)\\ &= \int\left(\frac{f(\boldsymbol{x},\boldsymbol{\theta}'+\boldsymbol{h}_i-\boldsymbol{h}_j)}{f(\boldsymbol{x},\boldsymbol{\theta}')}\right)^{1/2} f(\boldsymbol{x},\boldsymbol{\theta}')d\boldsymbol{x}d\boldsymbol{\theta}' \quad (19)\\ &= e^{\mu(1/2,\boldsymbol{h}_i-\boldsymbol{h}_j)} \quad (20)\end{aligned}$$

where a change of coordinates $\boldsymbol{\theta}' = \boldsymbol{\theta}+\boldsymbol{h}_j$ was performed. Analogous results can be obtained for the remaining expressions of the type (16). Substituting these into (8) yields (10).

When Θ' does not consist of the entire space $\mathbb{R}^m$, the reasoning above is not valid. The integration in (17) must be carried out only over those values of $\boldsymbol{x}$ and $\boldsymbol{\theta}$ for which $f(\boldsymbol{x},\boldsymbol{\theta}) > 0$, but this restriction is dropped in the transition to (18). If there exist values $(\boldsymbol{x},\boldsymbol{\theta})$ such that $f(\boldsymbol{x},\boldsymbol{\theta}) = 0$, $f(\boldsymbol{x},\boldsymbol{\theta}+\boldsymbol{h}_i) > 0$, and $f(\boldsymbol{x},\boldsymbol{\theta}+\boldsymbol{h}_j) > 0$, then those values are included in (18), but not in (17). This will always occur for some values of $\boldsymbol{h}_i$ and $\boldsymbol{h}_j$, unless $f(\boldsymbol{x},\boldsymbol{\theta}) > 0$ for all $\boldsymbol{\theta}$. Therefore, the value $E\{L^{1/2}(\boldsymbol{x};\boldsymbol{\theta}'+\boldsymbol{h}_i-\boldsymbol{h}_j,\boldsymbol{\theta}')\}$ can, in fact, be larger than $E\{L^{1/2}(\boldsymbol{x};\boldsymbol{\theta}+\boldsymbol{h}_i,\boldsymbol{\theta})L^{1/2}(\boldsymbol{x};\boldsymbol{\theta}+\boldsymbol{h}_j,\boldsymbol{\theta})\}$. We conclude that (10) does not necessarily hold unless Θ' includes the entire parameter space $\mathbb{R}^m$.

V. Counterexample

We now present an example which illustrates that (6) provides incorrect and even impossible results in some cases.

Consider the problem of estimating a scalar θ based on a single measurement x. Suppose that θ is uniformly distributed over the set $\Theta = [a,b] \cup [-b,-a]$, where $b > a > 0$ are given constants. Also suppose that the distribution of x conditioned on θ is Gaussian with mean θ and variance σ^2. The joint pdf of x and θ is thus given by

$$f(x,\theta) = \frac{e^{-(x-\theta)^2/2\sigma^2}}{2(b-a)\sqrt{2\pi\sigma^2}}\mathbb{1}_\Theta \quad (21)$$

where $\mathbb{1}_\Theta$ is an indicator function, which equals 1 when $\theta \in \Theta$ and 0 otherwise.

From (3), it follows that

$$e^{\mu(s,h)} = \frac{1}{2(b-a)}e^{-h^2s(1-s)/2\sigma^2}\int \mathbb{1}_\Theta\mathbb{1}_{\Theta+h}d\theta \quad (22)$$

where $\Theta+h = \{\theta+h : \theta\in\Theta\}$, and the integral equals the length of the intersection of the sets Θ and $\Theta+h$. On the other hand

$$\begin{aligned}\tilde{M}(s,h) &\triangleq E\left\{L^s(x;\theta+h,\theta)L^{1-s}(x;\theta-h,\theta)\right\} \\ &= \frac{1}{2(b-a)}e^{-2s(1-s)h^2/\sigma^2}\int\mathbb{1}_\Theta\mathbb{1}_{\Theta+h}\mathbb{1}_{\Theta-h}d\theta. \quad (23)\end{aligned}$$

Thus, contrary to (11), $e^{\mu(s,2h)}$ does not equal $\tilde{M}(s,h)$, since the latter depends on the length of the intersection of the three sets Θ, $\Theta+h$, and $\Theta-h$. Indeed, $\tilde{M}(s,h)$ is often substantially smaller than $e^{\mu(s,2h)}$, and as a result, use of (6) results in a "lower bound" which may exceed the true MSE.

[1]A slightly different version of this equation appears in [4, eq. (42)], the result of an obvious typographical error.

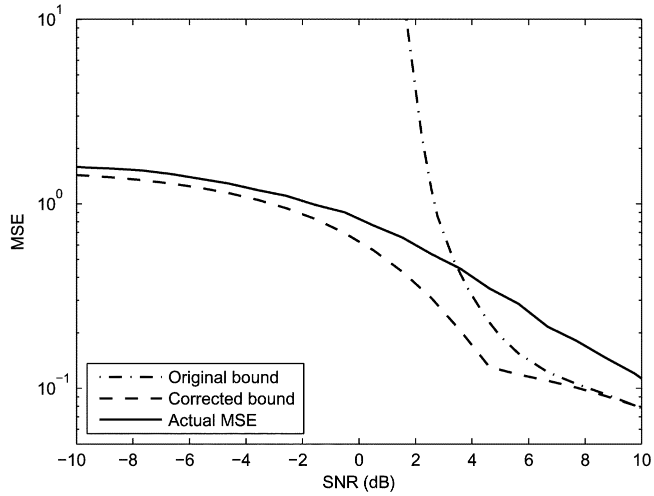

Fig. 1. Plot of the corrected WWB (24), the incorrect version (6), and the actual MSE, for a scalar estimation setting with a disjoint set Θ.

This problem is illustrated in Fig. 1, where the incorrect bound (6) is compared with the original WWB (4), which can be written as

$$\frac{h^2 e^{2\mu(s,h)}}{e^{\mu(2s,h)} + e^{\mu(2-2s,-h)} - 2\tilde{M}(s,h)}. \tag{24}$$

The actual MSE obtained by the optimal estimator can be calculated using Monte Carlo simulations, and is also plotted. In the figure, values of $a = 1/2$ and $b = 2$ were used. The variance σ^2 was modified to obtain various signal-to-noise ratios (SNRs), where $\mathrm{SNR} = \mathrm{Var}(\theta)/\sigma^2$.

It is evident from Fig. 1 that the value (6) becomes exceedingly high at low SNR. Indeed, for SNR values below approximately 0 dB, there always exist values of s and h such that the denominator of (6) is arbitrarily small, and thus the bound tends to infinity. For SNR values around 2–4 dB, (6) yields finite values which are larger than the actual MSE obtained by the optimal estimator. The original version (24), by contrast, closely follows the true MSE value.

Acknowledgment

The authors are grateful to Prof. A. J. Weiss and Prof. E. Weinstein for carefully reviewing an early version of this correspondence.

References

[1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968, vol. 1.

[2] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 3, pp. 386–391, May 1969.

[3] A. J. Weiss and E. Weinstein, "A lower bound on the mean-square error in random parameter estimation," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 5, pp. 680–682, Sep. 1985.

[4] E. Weinstein and A. J. Weiss, "A general class of lower bounds in parameter estimation," *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 338–342, Mar. 1988.

[5] T. J. Nohara and S. Haykin, "Application of the Weiss-Weinstein bound to a two-dimensional antenna array," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 9, pp. 1533–1534, Sep. 1988.

[6] D. F. DeLong, Use of the Weiss-Weinstein Bound to Compare the Direction-Finding Performance of Sparse Arrays Lincoln Lab, Mass. Inst. Tech, Cambridge, MA, 1993, Tech. Rep. 982.

[7] K. L. Bell, "Performance Bounds in Parameter Estimation With Application to Bearing Estimation," Ph.D. dissertation, George Mason University, Fairfax, VA, 1995.

Results of the Enumeration of Costas Arrays of Order 27

Konstantinos Drakakis, *Member, IEEE*,
Scott Rickard, *Senior Member, IEEE*,
James K. Beard, *Life Senior Member, IEEE*, Rodrigo Caballero,
Francesco Iorio, Gareth O'Brien, and John Walsh

***Abstract*—This correspondence presents the results of the enumeration of Costas arrays of order 27: all arrays found, except for one, are accounted for by the Golomb and Welch construction methods.**

***Index Terms*—Costas arrays, enumeration, Golomb method, order 27, Welch method.**

I. Introduction

In this brief note we present the results of the enumeration of Costas arrays of order 27. This result comes approximately 2.5 years after the last major enumeration project of Costas arrays undertaken, namely that for order 26, completed independently and by two different groups led by J. K. Beard [1] and S. Rickard [2], respectively. Our project was run on various supercomputers in Ireland [GridIreland[1], which actually ran 68.75% of the project, and some clusters in University College Dublin (Halation[2], Meteorite[3], Rowan)] and Scotland [the University of Edinburgh's EPCC's BlueGene[4]], as well as on several other private machines. Taking a CPU running at 2.00GHz as a reference, the project required approximately 25 years of single

Manuscript received May 23, 2008; revised June 16, 2008. Current version published September 17, 2008. This material was supported by the Science Foundation Ireland under Grant 05/YI2/I677.

K. Drakakis is with the School of Mathematics, University College Dublin, Belfield, Dublin 4, Ireland. He is also now with the Claude Shannon Institute and University College Dublin, Belfield, Dublin 4, Ireland (e-mail: konstantinos.drakakis@ucd.ie).

S. Rickard is with the School of Electrical, Electronic and Mechanical Engineering, University College Dublin, Belfield, Dublin 4, Ireland. He is also with the Claude Shannon Institute and University College Dublin, Belfield, Dublin 4, Ireland (e-mail: scott.rickard@ucd.ie).

J. K. Beard is an independent consultant (e-mail: jkbeard@ieee.org).

R. Caballero and G. O'Brien are with the School of Mathematics, University College Dublin, Belfield, Dublin 4, Ireland (e-mail: rodrigo.caballero@ucd.ie, gareth.obrien@ucd.ie).

F. Iorio is with the IBM Systems and Technology Group, IBM Technology Campus, Damastown Industrial Park, Mulhuddart, Dublin 15, Ireland (e-mail: francesco_iorio@ie.ibm.com).

J. Walsh is with the Department of Computer Science, Trinity College Dublin, College Green, Dublin 2, Ireland (e-mail: John.Walsh@cs.tcd.ie).

Communicated by G. Gong, Associate Editor for Sequences.

Color version Figure 1 in this correspondence are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2008.928979

[1]http://www.grid-ireland.org

[2]http://www.ucd.ie/itservices/researchit/services/clusters

[3]Internal UCD clusters

[4]http://www.epcc.ed.ac.uk/