

# Fast comparison of IS radar code sequences for lag profile inversion

M. S. Lehtinen<sup>1</sup>, I. I. Virtanen<sup>2</sup>, and J. Vierinen<sup>1</sup>

<sup>1</sup>Sodankylä Geophysical Observatory, 99600, Sodankylä, Finland

<sup>2</sup>Department of Physical Sciences, University of Oulu, P.O. Box 3000, 90014, Finland

Received: 3 January 2008 – Revised: 27 May 2008 – Accepted: 4 June 2008 – Published: 5 August 2008

**Abstract.** A fast method for theoretically comparing the posteriori variances produced by different phase code sequences in incoherent scatter radar (ISR) experiments is introduced. Alternating codes of types 1 and 2 are known to be optimal for selected range resolutions, but the code sets are inconveniently long for many purposes like ground clutter estimation and in cases where coherent echoes from lower ionospheric layers are to be analyzed in addition to standard F-layer spectra.

The method is used in practice for searching binary code quads that have estimation accuracy almost equal to that of much longer alternating code sets. Though the code sequences can consist of as few as four different transmission envelopes, the lag profile estimation variances are near to the theoretical minimum. Thus the short code sequence is equally good as a full cycle of alternating codes with the same pulse length and bit length. The short code groups cannot be directly decoded, but the decoding is done in connection with more computationally expensive lag profile inversion in data analysis.

The actual code searches as well as the analysis and real data results from the found short code searches are explained in other papers sent to the same issue of this journal. We also discuss interesting subtle differences found between the different alternating codes by this method. We assume that thermal noise dominates the incoherent scatter signal.

**Keywords.** Radio science (Ionospheric physics; Signal processing; Instruments and techniques)

## 1 Introduction

In incoherent scatter radar measurements the desired range resolution is often much better than the available pulse length

*Correspondence to:* M. S. Lehtinen  
(markku.lehtinen@sgo.fi)

of the radar. In order to maximize the statistics of the results special coding methods like alternating codes have been developed to allow the radar to be used at full duty cycle and still produce range resolutions much shorter than the pulse length used (Lehtinen and Häggstrom, 1987). Typical features of these kinds of methods are:

- Long pulses are modulated by a binary phase pattern, which gives range resolution approximately equal to the baud length of this pattern.
- This requires relatively little computation because the resolution is achieved by a simple summation/subtraction procedure of the corresponding lag products of all the codes in the code set.

Then there are the newer codes that have the advantage that the code set can contain fewer codes, especially important when the equivalent alternating code set has many codes (Virtanen et al., 2007; Dantie et al., 2002). These codes:

1. preserve the range resolution that the equivalent alternating code set provides,
2. but do not allow this resolution to be accessed by such a simple procedure as adding/subtracting corresponding lag products.
3. Instead they require a possibly computationally intensive inversion, and thus are practical now, but were not two decades ago.
4. Since these codes are found by a search technique, computationally intensive due to the vast number of possible code sets, it is necessary to make this process efficient.
5. The important thing is not whether data from a code group can be inverted, but rather how much the signal to noise ratio of the experiment is degraded by the inversion.

6. Since different code sets degrade the SNR by vastly different amounts, a very important part of the search technique is the evaluation of the degradation.
7. This paper describes a fast method for doing this, a method that is not necessarily as accurate as slower methods, but nonetheless is good enough to allow good code sets to be separated from the vast number of bad code sets.

We also use the developed comparison method to discuss some interesting differences found between strong alternating codes, weak alternating codes (Lehtinen and Häggstrom, 1987) and alternating codes of type 2 (Sulzer, 1993), seen in the way they perform in terms of variance behavior when post-integrating lag profile data to coarser resolutions.

The actual (accurate) method for decoding the codes is to apply linear statistical inversion to the lag profiles measurements, which represent the desired unknown plasma scattering lag profiles convolved with the range ambiguity functions corresponding to the different lagged products of the codes in the code group under study. If the lag profiles are very long, the faster method of inversion can be based on Fourier transforms of the profiles. These methods will fail near the ends of the profiles, but they provide us with a much faster way to estimate the variance behaviour of the inversion process with results valid if we do not worry about what happens near the ends of the profiles. Because the code groups are searched with time-consuming computer searches the speed obtained this way is essential.

The linear statistical inversion based method is described in Virtanen et al. (2007). These methods are developed towards general-purpose experiments in another paper in this special issue (Virtanen et al., 2008). There are also other related papers in this special issue:

Another paper (Vierinen et al., 2008) describes the efficient search algorithms used for actual searches of the code quads. We also discuss the advantages of using polyphase codes and amplitude modulation instead of binary modulation.

We discuss the possibility of doing inversion analysis of echo amplitude data to solve for the scattering amplitudes instead of correlation properties of the target (Vierinen et al., 2008b). This way of analysis is advantageous with very narrow targets like meteor heads and narrow ionospheric layers.

We have shown that the new experiments can be used together with advanced MCMC (Markov Chain Monte Carlo) techniques combined with ionospheric models to do analysis in cases beyond the capabilities of traditional GUISDAP analysis or the lag profile inversion (Kero et al., 2008).

We have also generalized binary alternating codes to a general number of phases (Markkanen et al., 2008). This paper is purely a generalization of the classical way of coding and decoding by summation – with no need for inversion methods.

## 2 Code selection

The focus of the code selection is to find the code sequences that produce minimum posteriori variances for the inverted lag profiles. Starting from the scattering process happening in the ionosphere it is possible to derive an analytical formula for the posteriori variance of any lag value for any code sequence. For each code length and each lag value the variance has a theoretical minimum. By comparing the calculated variances of different codes to the minimum variances at all lag values the best codes can be selected.

We use here a discrete formalism, where all times (and ranges which are also specified as corresponding times) are represented as integers. The discrete formalism can be derived from the continuous-time formalism developed in (Lehtinen, 1986; Lehtinen and Huuskonen, 1996) by studying samples of signals given by  $z_i = \frac{1}{\Delta t} \int_{i\Delta t}^{(i+1)\Delta t} z(t) dt$  in the limit  $\Delta t \rightarrow 0$ .

### 2.1 Scattering from the ionosphere

The scattering coefficient of an ionospheric plasma element is noted with  $x_{r,t}$ , where  $r$  is signal travel time to the scattering volume and back and  $t$  is time. The scattering is modeled as a zero mean stochastic process with covariance

$$\langle x_{r,t} \overline{x_{r',t'}} \rangle = \delta_{rr'} \sigma_r^{t-t'}, \quad (1)$$

where  $\delta_{rr'}$  is the Kronecker delta and  $\bar{x}$  is the complex conjugate of  $x$ . Plasma scattering properties are modeled by the plasma autocorrelation function  $\sigma_r^{t-t'}$ .

In a practical incoherent scatter experiment the illumination by radio waves is not continuous, but the radar sends phase coded pulses to the ionosphere. The scattered signal is recorded in the form of discrete IQ-samples (raw data).

If the discretization step  $\Delta t$  is small enough, the amplitude ambiguity function for range simplifies to be equal to the transmission waveform  $\text{env}_t$  and thus the received complex signal is a weighted sum of the scattering coefficients

$$z_q = \sum_{r=-\infty}^{\infty} \text{env}_r x_{q-r;q} + \varepsilon_q, \quad (2)$$

where  $\text{env}_q$  is the complex transmission envelope and  $\varepsilon_q$  the measurement error due to thermal noise. Its correlation properties are expressed by

$$\langle \varepsilon_q \overline{\varepsilon_{q'}} \rangle = \frac{C}{\Delta t} \delta_{qq'}, \quad (3)$$

where  $C$  is a constant depending on the amplitude of the thermal noise.

### 2.2 Lag profiles and inversion

In an incoherent scatter radar experiment lagged products of the received signal and its complex conjugate,

$$m_{q;\tau} = z_q \overline{z_{q-\tau}}, \quad (4)$$

are calculated. Using Eq. (1) the expectations of the lagged products can be written in the form

$$\begin{aligned} \langle m_{q;\tau} \rangle &= \langle z_q \overline{z_{q-\tau}} \rangle \\ &= \left\langle \sum_{r=-\infty}^{\infty} (\text{env}_{r;\tau} x_{q-r;q} + \varepsilon_{q;\tau}) \cdot \overline{\sum_{s=-\infty}^{\infty} (\text{env}_{s;\tau} x_{q-\tau-s;q-\tau} + \varepsilon_{q-\tau})} \right\rangle \\ &= \sum_{r=-\infty}^{\infty} \text{env}_r \overline{\text{env}_{r-\tau}} \sigma_{q-r}^\tau \\ &= \sum_{r=-\infty}^{\infty} W_{r;\tau} \sigma_{q-r}^\tau. \end{aligned} \tag{5}$$

The expectations of the terms containing measurement errors are zeros.

A range profile of plasma autocorrelation function  $\sigma_q^\tau$  with a fixed value of time lag  $\tau$  is called a ‘‘lag profile’’. In the analysis several lag profiles  $\sigma_q^\tau$  with selected values of  $\tau$  are solved as separate problems to sample the whole plasma autocorrelation function. The difference between a measured ambiguous lagged product and its expectation is

$$\begin{aligned} m_{q;\tau} &= \langle m_{q;\tau} \rangle + \varepsilon_{q;\tau} \\ &= \sum_{r=-\infty}^{\infty} W_{r;\tau} \sigma_{q-r}^\tau + \varepsilon_{q;\tau}. \end{aligned} \tag{6}$$

The error term in Eq. (6) is rather complicated. The fourth moment’s theorem for complex Gaussian processes imply

$$\begin{aligned} \langle \varepsilon_{q;\tau} \overline{\varepsilon_{q';\tau'}} \rangle &= \langle z_q \overline{z_{q'-\tau'}} \rangle \langle z_{q'} \overline{z_{q-\tau}} \rangle \\ \langle \varepsilon_{q;\tau} \varepsilon_{q';\tau'} \rangle &= \langle z_q \overline{z_{q'}} \rangle \langle \overline{z_{q-\tau}} z_{q'-\tau'} \rangle. \end{aligned} \tag{7}$$

These formulas, represented in terms of the real and imaginary parts of the lag profile expectations  $\langle m_{q;\tau} \rangle$  can be found in the references cited above. The mutual correlations of the error term  $\varepsilon_{q;\tau}$  between both different lags and different ranges can be fully explained by these formulas. This is the correct way to model the self-noise of the lag profile estimates, when the signal-to-noise ratio is not very low.

For the purposes of this study, we assume a very low signal-to-noise ratio, and the error term simplifies to

$$\begin{aligned} \langle \text{Re}(\varepsilon_{q;\tau}) \text{Re}(\varepsilon_{q';\tau'}) \rangle &= \frac{C^2}{\Delta t^2} \delta_{qq'} (\delta_{\tau\tau'} + \delta_{\tau 0}) \\ \langle \text{Im}(\varepsilon_{q;\tau}) \text{Im}(\varepsilon_{q';\tau'}) \rangle &= \frac{C^2}{\Delta t^2} \delta_{qq'} \delta_{\tau\tau'} (1 - \delta_{\tau 0}) \\ \langle \text{Re}(\varepsilon_{q;\tau}) \text{Im}(\varepsilon_{q';\tau'}) \rangle &= 0, \end{aligned} \tag{8}$$

if  $\tau \geq 0$  and  $\tau' \geq 0$ . The lag profile estimates are thus independent of each other if either the lag values or the range values differ. The variance of the zero lag profile is two times the variance of the other lag profiles. Negative values of the lags are completely dependent on the positive values, being complex conjugates of each other.

The Fourier transform of the measured lag profile is

$$\begin{aligned} \tilde{m}_\tau(\omega) &= \sum_{q=-\infty}^{\infty} e^{iq\omega} \left( \sum_{r=-\infty}^{\infty} W_{r;\tau} \sigma_{q-r}^\tau + \varepsilon_{q;\tau} \right) \\ &= \tilde{\sigma}^\tau(\omega) \tilde{W}_\tau(\omega) + \tilde{\varepsilon}_\tau(\omega) \end{aligned} \tag{9}$$

as indicated by the convolution theorem of the Fourier transforms. The Fourier transform of the plasma autocorrelation function  $\sigma_r^\tau$  is not the plasma scattering spectrum, because the transform is taken with respect to the range variable  $r$ , not the time lag  $\tau$ .

The inversion solution (and the error term) of the Fourier transform of the unknown lag profile is

$$\tilde{\sigma}^\tau(\omega) = \frac{\tilde{m}_\tau(\omega)}{\tilde{W}_\tau(\omega)} - \frac{\tilde{\varepsilon}_\tau(\omega)}{\tilde{W}_\tau(\omega)}. \tag{10}$$

### 2.3 Posteriori noise

Now lets take a closer look at the transform of the noise, i.e. the last term in Eq. (9). Using Eq. (8) its covariance can be written as

$$\begin{aligned} \langle \tilde{\varepsilon}_\tau(\omega) \overline{\tilde{\varepsilon}_\tau(\omega')} \rangle &= \sum_{q=-\infty}^{\infty} e^{iq(\omega-\omega')} \\ &= 2\pi \delta(\omega - \omega'). \end{aligned} \tag{11}$$

The latter term in Eq. (10) is the Fourier transform of the posteriori noise. By taking the inverse transform we get the posteriori noise in time domain

$$\xi = \mathcal{F}^{-1}(\tilde{\varepsilon}_\tau(\omega) / \tilde{W}_\tau(\omega)). \tag{12}$$

The covariance of the posteriori noise in Eq. (12) is

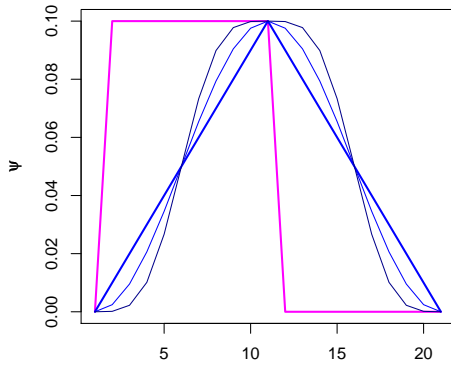
$$\begin{aligned} \langle \xi_p \overline{\xi_q} \rangle &= \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} e^{-ip\omega} e^{iq\omega'} \left\langle \frac{\tilde{\varepsilon}_\tau(\omega)}{\tilde{W}_\tau(\omega)} \overline{\frac{\tilde{\varepsilon}_\tau(\omega')}{\tilde{W}_\tau(\omega')}} \right\rangle d\omega d\omega' \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{-i((p-q)\omega)} \frac{1}{|\tilde{W}_\tau(\omega)|^2} d\omega. \end{aligned} \tag{13}$$

The previous holds for a single transmission envelope that leads to a single range ambiguity function for each lag value. In case of several envelopes, i.e. a code sequence longer than one, each of the envelopes produces a different range ambiguity function for each lag value. From a sequence of  $N$  codes, equations

$$W_{r;\tau}^u = \text{env}_r^u \overline{\text{env}_{r-\tau}^u}, \tag{14}$$

$u = 1 \dots N$ , are produced. Each of these range ambiguity functions produce an equation similar to Eq. (6)

$$m_{q;\tau}^u = \sum_{r=-\infty}^{\infty} W_{r;\tau}^u \sigma_{q-r}^\tau + \varepsilon_{q;\tau}^u \tag{15}$$



**Fig. 1.** Range integration averaging kernels used in the calculations. The boxcar kernel (magenta) was not used in the subsequent calculations. The triangle kernel (thick blue) seemed to perform best while the sin-shaped kernel (thin blue) and the iterated sin-shaped kernel (dark blue) performed slightly worse. The results corresponding to the last three kernels are shown with correspondingly colored curves in the results.

and has a Fourier transform

$$\tilde{m}_\tau^u(\omega) = \tilde{\sigma}^\tau(\omega) \tilde{W}_\tau^u(\omega) + \tilde{\varepsilon}_\tau^u(\omega). \quad (16)$$

The inversion solution (and the error term) for the Fourier transform of the lag profile using all the measurements can now be written as

$$\tilde{\sigma}_\tau(\omega) = \frac{\sum_{u=1}^N \overline{\tilde{W}_\tau^u(\omega) \tilde{m}_\tau^u(\omega)}}{\sum_{u=1}^N |\tilde{W}_\tau^u(\omega)|^2} + \frac{\tilde{\varepsilon}_\tau(\omega)}{\sqrt{\sum_{u=1}^N |\tilde{W}_\tau^u(\omega)|^2}}, \quad (17)$$

where  $\tilde{\varepsilon}_\tau(\omega)$  is the Fourier transform of a normalized white noise. The posteriori noise covariance is

$$\langle \xi_p \overline{\xi_q} \rangle = \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{-i(p-q)\omega}}{\sum_{u=1}^N |\tilde{W}_\tau^u(\omega)|^2} d\omega. \quad (18)$$

From Eq. (18) we get the posteriori variance of the inversion result

$$\langle \xi_p \overline{\xi_p} \rangle = \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{\sum_{u=1}^N |\tilde{W}_\tau^u(\omega)|^2} d\omega. \quad (19)$$

that has a minimum

$$\langle \xi_p \overline{\xi_p} \rangle \geq \left( \sum_{u=1}^N \sum_{r=-\infty}^{\infty} |W_{r;\tau}^u|^2 \right)^{-1}. \quad (20)$$

### 2.4 Proof of the accuracy limit

It is surprisingly easy to see that the inequality (20) must be true. For this, let us assume that the only unknown lag profile value is  $\sigma_q^\tau$  and that we know all the other values of  $\sigma_p^\tau$ , if

$p \neq q$ . Our only unknown then appears in the measurement Eq. (15) multiplied by all the values  $W_{r;\tau}^u$  with  $1 \leq u \leq N$  and  $-\infty < r < \infty$ .

As the error terms in Eq. (20) are assumed normalized and independent, the inversion accuracy for this only unknown variable is easily seen to be equal to the right hand side of Eq. (20).

If we now drop the assumption about exact knowledge of the other scattering terms  $\sigma_p^\tau$ ,  $p \neq q$ , the estimation accuracy of the previously only unknown cannot get any better. Thus, the inequality in Eq. (20) must be true.

The inequality (20) is also a consequence of the well-known mathematical fact that the harmonic mean of numbers is smaller or equal to the arithmetic mean of the numbers.

### 2.5 Code selection criteria

Now we have finally arrived to the formulas that can be used in comparing different code sequences. The efficiency of a code is defined as ratio of the minimum posteriori variance as defined in Eq. (20) and the actual posteriori variance of the code under study,

$$R = \frac{\left( \sum_{u=1}^N \sum_{r=-\infty}^{\infty} |W_{r;\tau}^u|^2 \right)^{-1}}{\frac{1}{2\pi} \int_0^{2\pi} \frac{1}{\sum_{u=1}^N |\tilde{W}_\tau^u(\omega)|^2} d\omega}. \quad (21)$$

In practical calculations the Fourier transform of the range ambiguity function can be handled with DFT. When the integral in the denominator is written in discrete form using Riemann sum, we get the form

$$\frac{\left( \sum_{u=1}^N \sum_{r=-\infty}^{\infty} |W_{r;\tau}^u|^2 \right)^{-1}}{\frac{1}{M} \sum_{q=0}^{M-1} \frac{1}{\sum_{u=1}^N |\tilde{W}_{\tau,q}^u|^2}} \rightarrow R \text{ as } M \rightarrow \infty. \quad (22)$$

In the case of binary code groups the nominator can be simplified even further. The amplitude of range ambiguity function is  $|W_{r;\tau}^u|=1$  in all non-zero parts of the function. If the codes have  $N_b$  bits and only the full lags (multiples of the bit length) are studied, Eq. (22) can be written in form

$$\frac{(N(N_b - \tau))^{-1}}{\frac{1}{M} \sum_{q=0}^{M-1} \frac{1}{\sum_{u=1}^N |\tilde{W}_{\tau,q}^u|^2}} \rightarrow R \text{ as } M \rightarrow \infty. \quad (23)$$

### 2.6 Range integration

As lag profile inversion is usually not performed with the best possible range resolution, but range integration is used at least at larger ranges, it is useful to investigate how the code sequences behave in the integration process. Whereas the maximum range resolution is achieved by using the original range ambiguity samples in the inversion process, convolutions of range ambiguity function and an averaging function

$\psi$  are used to get the range integrated values. The convolution gives an extra term to the posteriori variance (Eq. 19), so that the variance of the integrated lag profile is

$$\langle \xi_p \overline{\xi_p} \rangle = \frac{1}{2\pi} \int_0^{2\pi} \frac{|\tilde{\Psi}(\omega)|^2}{\sum_{u=1}^N |\tilde{W}_\tau^u(\omega)|^2} d\omega, \quad (24)$$

where  $\tilde{\Psi}(\omega)$  is the Fourier transform of the averaging convolution kernel  $\Psi$ . The discrete version of this formula is easily obtained by replacing  $1/2\pi$  by  $1/M$  and the integral over  $\omega$  by summation over the discrete FFT values  $\omega_i$ . A very natural choice for the averaging kernel would be the boxcar function

$$\Psi_{\text{box}}(n) = \frac{1}{N_r}, \text{ if } 0 \leq n < N_r, \quad \Psi(n) = 0 \text{ otherwise.} \quad (25)$$

However, it was discovered that the boxcar shape performed very badly. For this reason we defined a triangular averaging function by setting

$$\begin{aligned} \Delta_{N_r}(n) &= n/N_r, \text{ if } 1 \leq n \leq N_r \\ \Delta_{N_r}(n) &= 1 - n/N_r, \text{ if } N_r < n \leq 2N_r \\ \Delta_{N_r}(n) &= 0, \text{ otherwise} \end{aligned} \quad (26)$$

and

$$\Psi_{\Delta}(n) = \Delta_{N_r}/N_r. \quad (27)$$

Moreover, a sinusoidal shaped averaging function as well as an iterated sinusoidal averaging function (rather close to boxcar) were defined by

$$\Psi_{\text{sin}}(n) = \sin\left(\frac{\pi}{2} \Delta_{N_r}(n)\right)^2 \quad (28)$$

and

$$\Psi_{\text{sin sin}}(n) = \sin\left(\frac{\pi}{2} \sin\left(\frac{\pi}{2} \Delta_{N_r}(n)\right)^2\right)^2. \quad (29)$$

All these kernels are shown in Fig. 1.

### 3 Numerical results

As our goal was to find a code sequence of 15 bauds of  $10 \mu\text{s}$  duration, we first evaluated such a code set of randomly chosen binary bauds. The results for the randomly chosen code set are shown in Fig. 3. We have plotted the variances of the experiments as a function of range integration. The results from the simple and fast formula (22) or (23) are shown by the red curve and the results for more careful calculations, modelling the signal values as 5 samples taken for each baud by the blue curves. The three different rows of the curves correspond to a different number of fractional lag profiles taken to represent each full lag value.

The formulas used for both cases are the same. The main difference is that in the simple case we use 1 discrete sample

to represent each baud, while in the more realistic case we use 5 samples to represent each baud. Moreover, in the more realistic case the full lags are multiples of 5 samples and each of them is estimated using fractional lags around this exact multiple of 5 samples.

It can be seen that realistic curves rather well follow the fast formula (22) or (23), but that there are occasionally even significant discrepancies between them. However, we feel this anyway justifies the use of rapid formulae in extensive computer searches.

### 4 Careful explanation of the result figures

The main contribution of this paper, in addition to Eqs. (22) or (23) and (24) are the numerical results shown in form of matrices of curves for different situations. As these graphics contain a great wealth of information, it is necessary to carefully explain everything shown in them.

The figures consist of three of five rows of panes. The first row shows the pulse forms and the range ambiguity functions. The second row shows the squared moduli of the Fourier transforms of the individual range ambiguities, normalized by dividing by the number of codes and the sum of these individual Fourier transforms. This sum can be considered as the average Fisher information of the inverse convolution problem. The last 1 or 3 rows show how the results behave when range-integrated to various resolutions. We show the behaviour for both the simplest formula as the red curve as well as a set of other curves – calculated more carefully by discretizing the analysis to 5 steps for each baud in the code. We always show the results for calculation of fractional lags  $\pm 2, \pm 1$  and 0 and in the 5-row panels also for other kinds of fractional lag sets involved in the analysis.

#### 4.1 First line of figure panes

Here we simply show the range ambiguity functions for all full lags. LAG 0 is different, as the range ambiguity function is just a constant of the length of the pulse itself and it is not very informative to show it. Instead, we plot the codes themselves in that position. To emphasize the difference, we plot the code in blue. At the top of these panels we also print the relevant full lag number.

#### 4.2 Second line of figure panes

On the second line of figure panes we plot the average of the squared moduli of the Fourier transforms of the ambiguity functions, calculated by simple (no fractional samples) formulas. As the convolution inverse problem diagonalizes on the Fourier side, this curve can be understood as the representation of the diagonal Fisher information matrix of the inverse problem of deconvolving the ambiguity functions. This function appears in the denominator of the integral in Eq. (19).

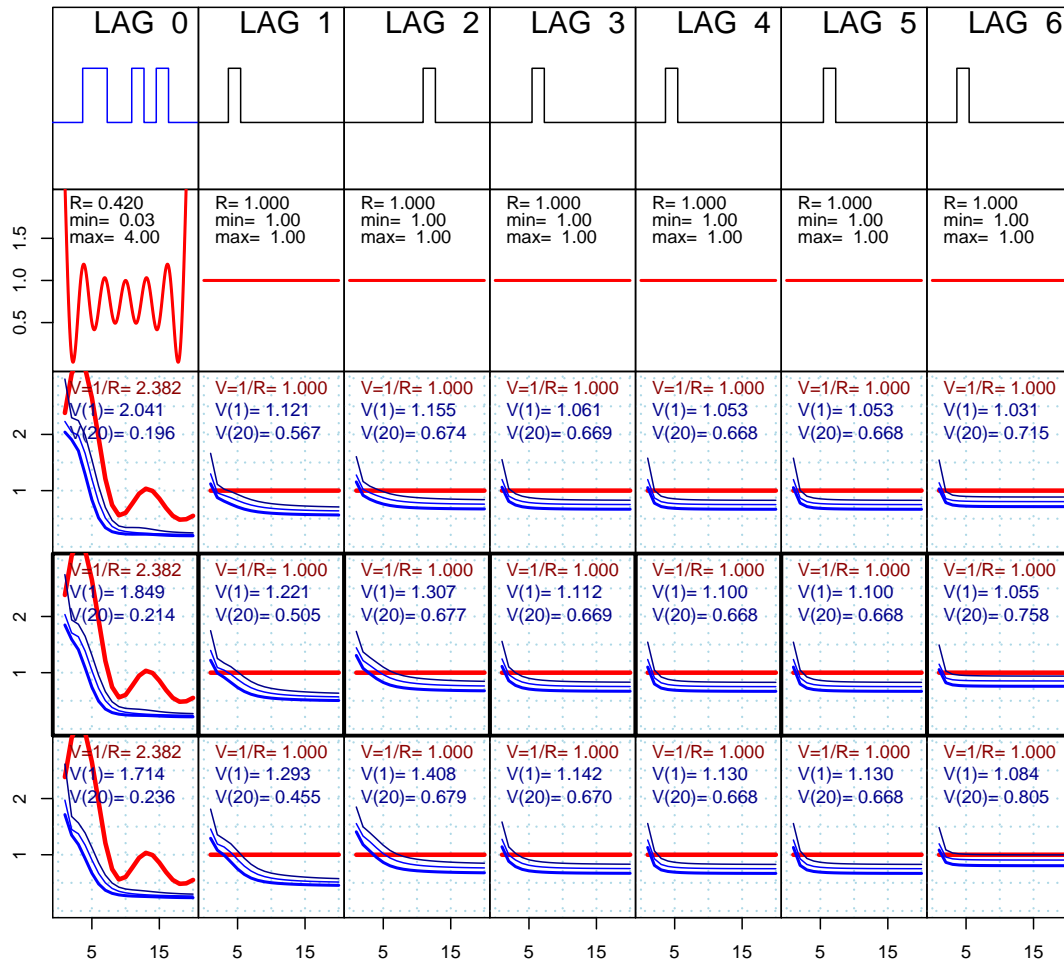


Fig. 2. Evaluation of a simple multiple-pulse type code.

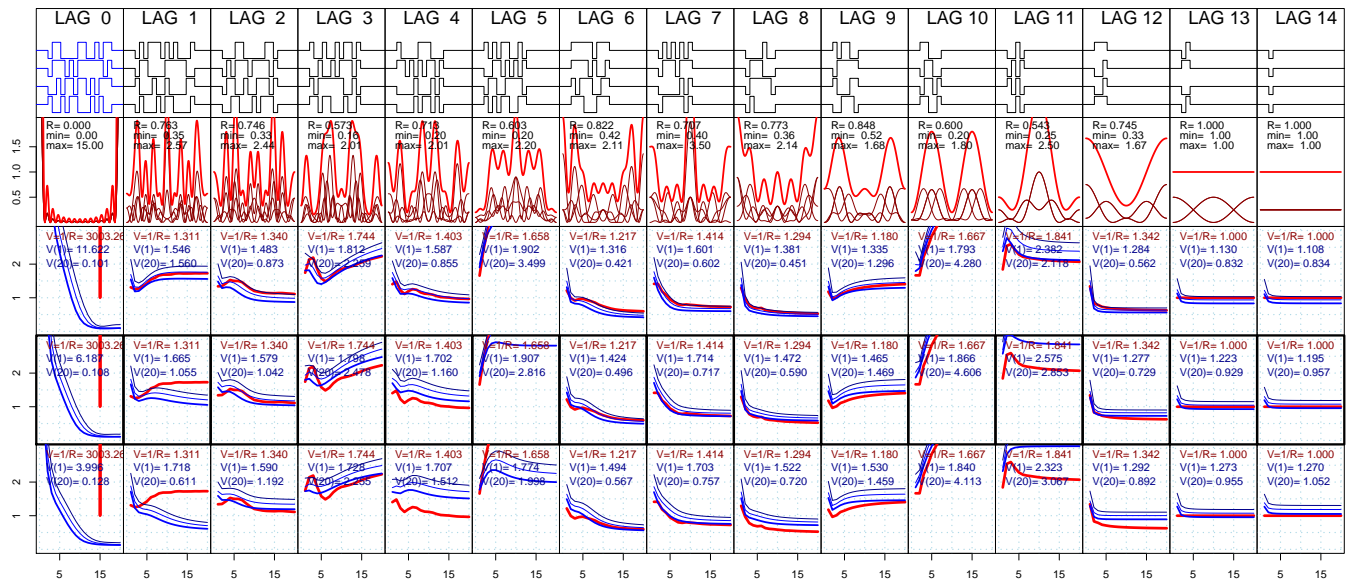
If the ambiguity is a Kronecker delta of unit magnitude, the corresponding Fisher information is a constant one in Fourier side. Thus, a constant one represents a set of codes representing exactly the same kind of information that can be obtained exactly from a single unit-strength pulse pair. A constant unit value here represents a perfect code set in the sense that the statistical properties of the analysis result are exactly equal to those corresponding to a simple pulse pair. However, we have normalized this curve by dividing it by the length of the ambiguity function – so it actually means that a constant unit value means pulse compression perfectly done with an apparent power corresponding to the length of the ambiguity function.

Please, note that the frequency axis is that of the FFT software: zero frequency corresponds to the left side of the figure while small negative frequencies are found on the right side of the x-axis. The center of the x-axis corresponds to the highest frequencies seen with the time discretization chosen.

The Fourier side information is easily interpreted in terms of what happens with range integration. If we have much information at zero or low frequencies, it is obvious that range integration should perform well. This happens clearly in lag 8 in Fig. 4. Also, zero lags integrate well if fractional lags are taken into account. For the simple discrete case zero lags are singular and integration cannot help, except for the case of integration length equal to pulse length, which is seen as the downward spike on the red curve in some panels. On the other hand, having little information at DC results in very bad range integration results. This is remarkably clear for lags 5 and 10 in Fig. 3.

#### 4.3 Last rows of figure panes

The last 1 row or 3 rows of the figure panes show the behaviour of the codes, when range-integrated to resolution between 1 and 20 baud lengths. If we show 3 rows, the difference between those is in the number of fractional lags we include to form the final estimate for each full lag. The



**Fig. 3.** Evaluation of a randomly chosen 4-code binary sequence of length 15. In the first row of panels the ambiguity functions for each lag are shown, corresponding to all full lag values. For lag 0 we show the codes themselves (drawn in blue), as the ambiguity functions would be just constants shapes of length equal to the length of the total pulse. The second row of panels shows the average of the squared moduli of the Fourier transforms of the ambiguity over the code set (red) and also the individual terms in this average shown by the darker red lines in bottom. The lower three panels show the behavior of variances of lag profiles if integrated to different resolutions between 1 and 20 baud lengths. The fourth row corresponds to using fractional lags of  $\pm 2$ ,  $\pm 1$  and 0 to represent each full lag, while in the third row we have used fractional lags  $\pm 1$  and 0 and in the fifth row fractional lags  $\pm 3$ ,  $\pm 2$ ,  $\pm 1$  and 0. The red curves in all the rows correspond to a simple calculation using no fractional lags at all, while the blue curves correspond to more realistic analysis, all using the fractional lags specified but 3 different kernels for range integration. The simple red curves are used for fast computer searches and we see that they rather well approximate the more realistic variance estimates.

alternatives are fractional lags  $\pm 1$  and 0 (3rd row),  $\pm 2 \pm 1$  and 0 (4th row) and  $\pm 3 \pm 2 \pm 1$  and 0 (5th row). If only one row is shown, the fractional lags are  $\pm 2 \pm 1$  and 0, which is the most natural choice, as for 5 samples for each baud this results in each fractional lag profile being used once and only once for each full lag profile estimation.

In each of these figures we have plotted in red the simplest possible result for range integration with only full lags taken into account in Eq. (24). This is to make it easy to compare the simple formula to the more careful variance calculations shown by the thick blue, thin blue and dark blue curves corresponding to more careful calculations, discretized at 1/5 of the baud length and using the different range integration kernels defined in Sect. 2.6.

#### 4.4 Normalization of the curves

It is important to understand that we have normalized all the variance curves by dividing them with the number of independent estimates going into the analysis. These factors include:

- Number of codes in the sequence
- Number of fractional lags

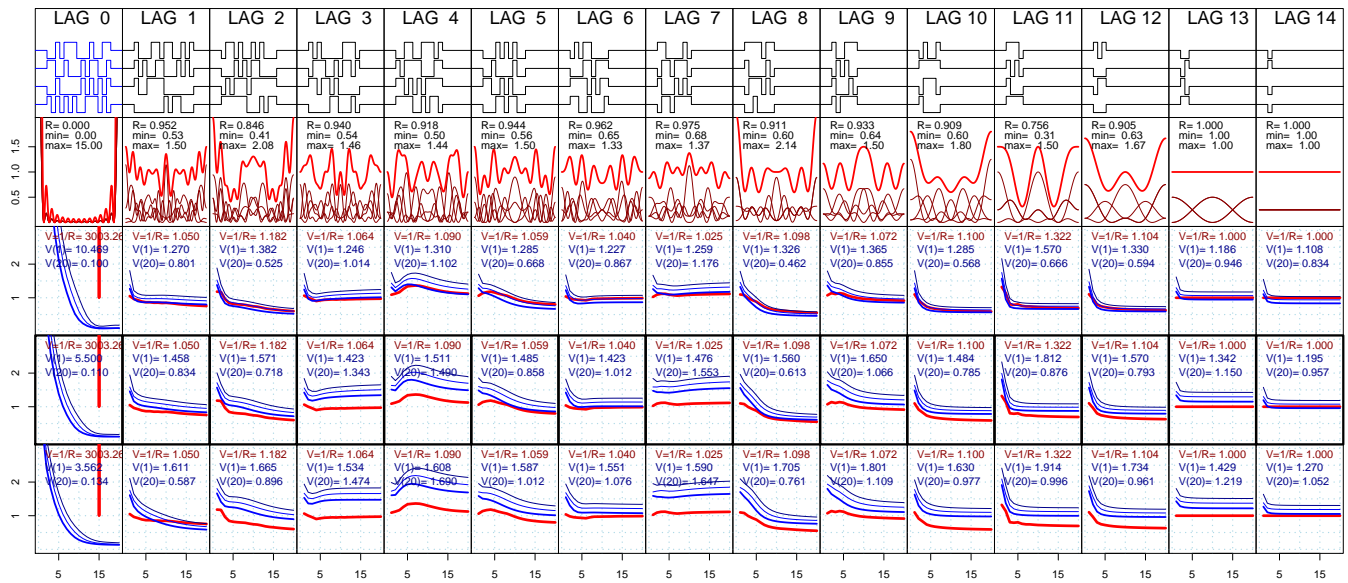
#### – Range integration

Also, the constant  $C$  in Eq. (3) is chosen to be equal to the baud length, so that for the simplest situations – discretized to the baud length – the signal thermal noise is assumed to be of unit variance. In the more accurate calculations discretizing the signal so that each baud corresponds to  $N_{\text{flag}}$  samples, the signal variances are equal to  $N_{\text{flag}}$  as it should be for sampled white noise.

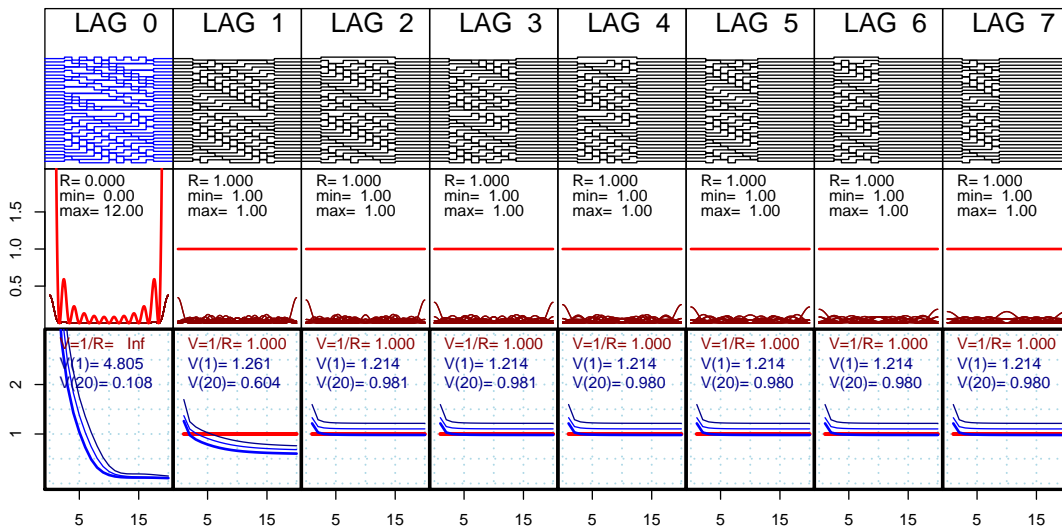
The reason for these choices is to make the scales of the different curves comparable. All variances shown in the curves should actually be equal to 1 if they would be calculated by simple estimates just taking into account the number of independent data going to the analysis. The difference from 1 is then a measure of non-perfectness of the inversion decoding process and the range-post-integration process caused by correlations of the errors of the inversion process as well as detailed behaviour of the ambiguity functions and the post-integration kernel.

Without this normalization the values of the variance curves at range-integration 20 should be 1/20 of the value they have at range-integration 1, which would make these curves difficult to read. This way the reader may understand the curves as additional correction factors, which must be





**Fig. 4.** Evaluation of the chosen code for a 15-baud 4-pulse experiment. While not perfect, the variances as function of range integration perform significantly better than in the previous figure with a randomly chosen code. For lags 2, 8, 10, 11 and 12 the variances go down faster than the range integration would imply. This is connected with the fact that the Fourier side information curve (red curve in the second row of panels) shows high values for DC and low frequencies.



**Fig. 5.** An alternating code set made from a 16-bit strong alternating code by truncating to 12 bits. We have only shown the most natural fractional lag set:  $\pm 2/5$ ,  $\pm 1/5$  and 0. It is clear that all lags  $\geq 2$  display a perfect decoding behaviour – they behave as if plain 2-pulse pairs corresponding to all 2-pulse-pairs in the set would have been available.

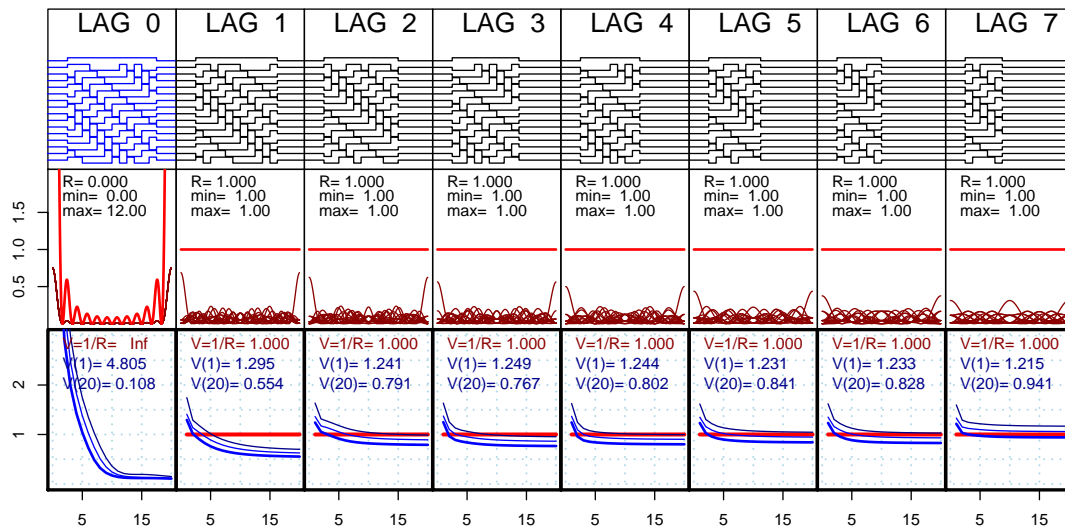
included after simple scalings taking into account the above listed factors are used.

#### 4.5 Numerical data shown in the panes

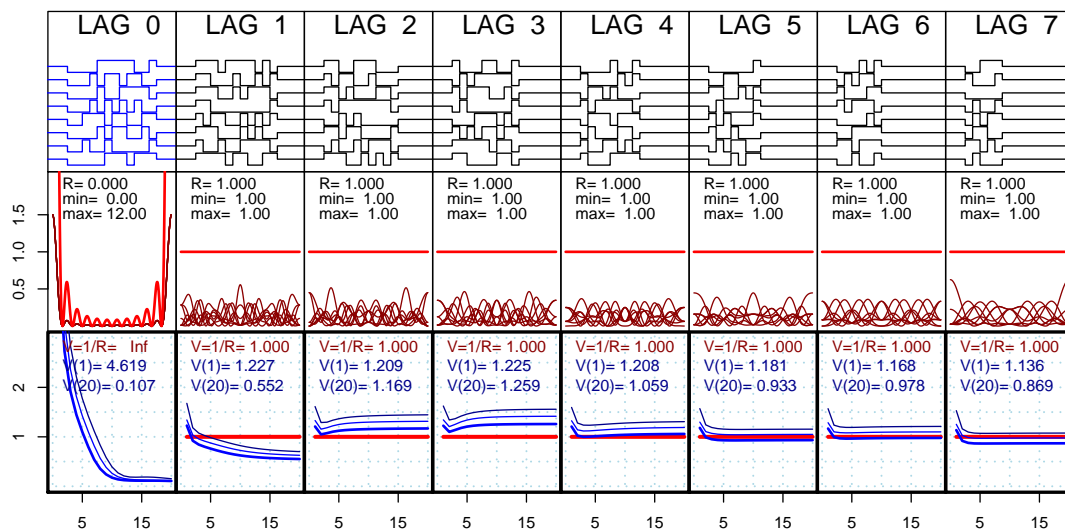
In the panes we have also helped the inspector by showing some numeric data relevant to the curves. We have shown

the  $R$  value defined in (Lehtinen et al., 2004) and (Damtie et al., 2002) to describe the close-to-perfectness of a code (or here a code set) on the panels of the second row for each lag profile. There we also print the minimum and maximum values of the Fourier side information curve, to make it easier to see how much it differs from a constant 1.





**Fig. 6.** A 12-bit weak alternating code set. Some differences to the strong set are visible. While slightly (but insignificantly) worse than the strong set in case of no range integration, in some cases the integration behaviour is actually better.



**Fig. 7.** Alternating code set of length 12 and type 2 introduced by Sulzer. Again the coarse analysis (red curve) represents perfect behaviour, but it seems that in integration and realistically accurate fractional lag calculations the first lags seem to behave somewhat worse than the alternating codes do.

The inverse of  $R$ ,  $V=1/R$  corresponds to the variance of a normalized situation and we have shown that value on the last rows of panel plotted. The corresponding variance calculated more accurately is also shown in dark blue on the same panels, as well as the normalized variance integrated to 1 baud length –  $V(1)$  – or to 20 baud lengths –  $V(20)$ .

### 5 Interesting behaviour of different alternating codes

Alternating codes of various types show rather interesting small differences in behaviour. The reason for this is that while the perfect decoding properties exist for all pulse pairs in alternating codes, essentially only full lags are considered and finite baud lengths are ignored. Strong alternating codes eliminate some part of this interplay by guaranteeing disappearance of unwanted responses also for slightly offset lags, but the other alternating code versions are not equally good in this respect.

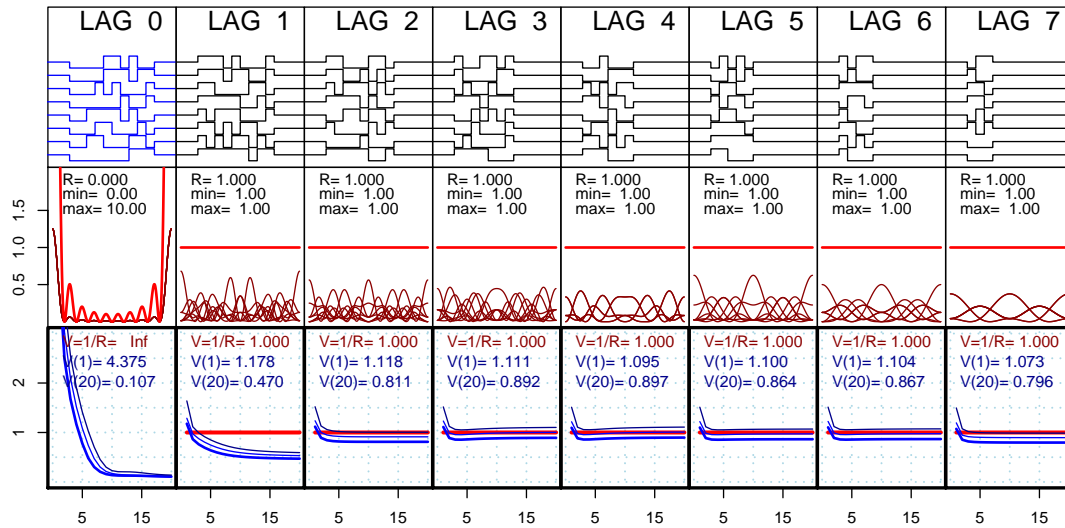


Fig. 8. Alternating code set of length 10 and type 2. In this set the problem with integration seen in the set of length 12 (previous picture) is not seen.

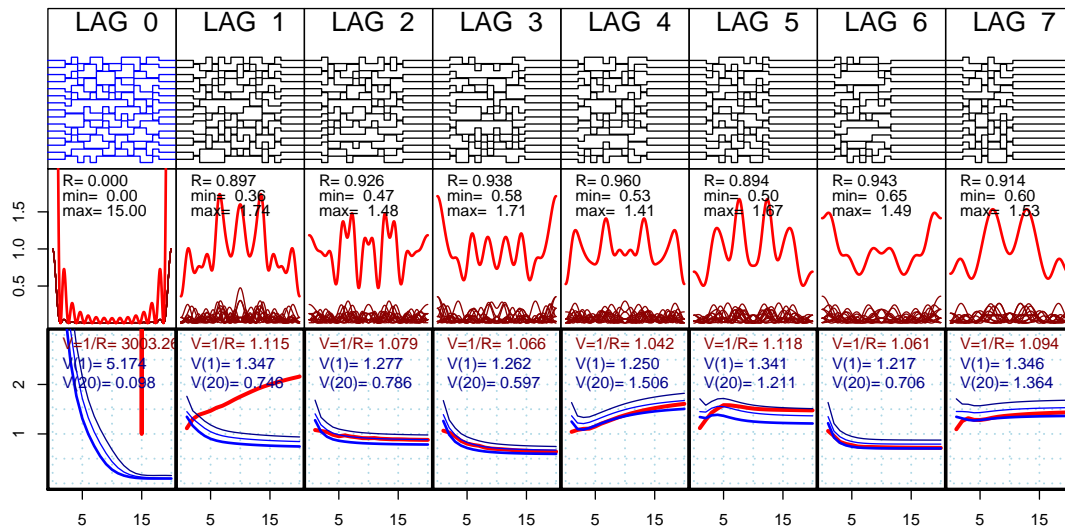


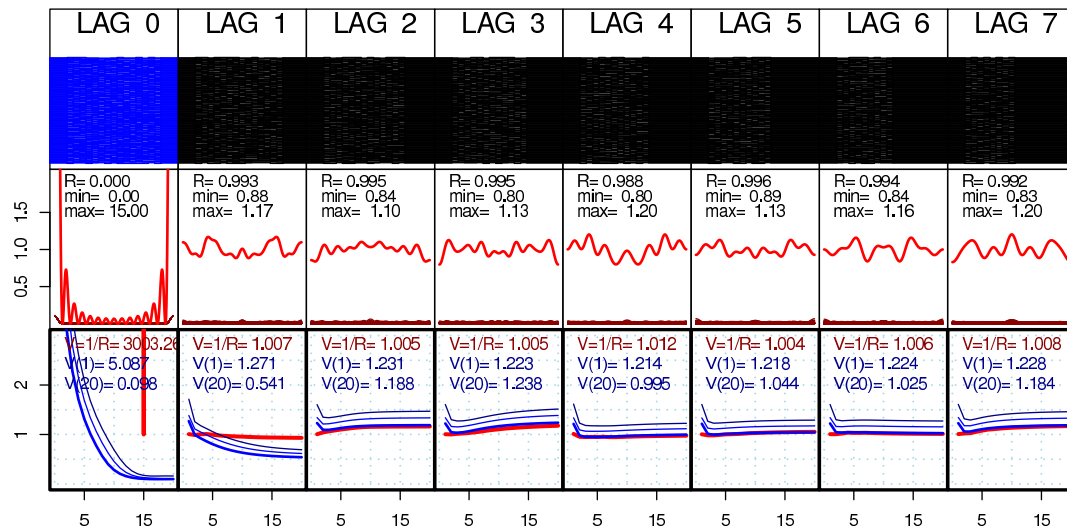
Fig. 9. A set of 15 randomly chosen pulses of 15 bauds each. It seems to be roughly equally good as the chosen set of 4 pulses in Fig. 4.

Very interestingly, weak alternating codes (Fig. 6) may actually perform better in range integration than strong alternating codes (Fig. 5). This is because they probably have less phase changes in the ambiguity functions than the strong ones. They are less perfectly similar to independent pulse pairs, and this can actually help regarding integration. For example: a pair of pulses with no phase shift between them behaves better in integration than a single pulse, while of course worse regarding the original resolution.

Type 2 alternating codes of length 12 also show a degradation of performance in range integration (Fig. 7), while those codes of length 10 do not seem to show this (Fig. 8). This

is due to the detailed working of amplitude functions in the sub-baud range resolution, but what exactly happens cannot be explained here.

We have also evaluated the behaviour of the first 7 lags of randomly chosen code sets of lengths 15 and 150, respectively (Figs. 9 and 10). These figures show that neither of these sets seems to behave as well as any of the alternating code sets. Compared to the set of 4 specially chosen codes (Fig. 4), the set of 15 randomly chosen pulses seems to be of approximately equal performance.



**Fig. 10.** A set of 150 randomly chosen pulses of 15 bauds each. This many pulses are necessary to make random code sets to be comparable to alternating code sets.

## 6 Conclusions

We have developed a fast method useful in comparing variances of lag profiles and range-integrated lag profiles using sets of codes in incoherent scatter radar measurements. We have shown that faster and more approximative ways to apply the method produce approximatively the same kinds of results as better modeled and more realistic ways to apply the method. This makes it possible to use the method for extensive computer searches of codes, where speed is essential.

We have also shown that the behavior of a new 4-pulse 15-baud code (Fig. 4) is almost as good as known methods of same number of bauds but with much larger number of pulses in the cycle.

In addition, we have found subtle differences in the behavior of the previously known different types of alternating codes and have seen that randomly chosen code sets need to be very long to produce equally good results as the other methods discussed.

*Acknowledgements.* This work was supported by the Academy of Finland (application number 213476, Finnish Programme for Centres of Excellence in Research 2006–2011; application number 43988, EISCAT Data Analysis and Research and by the Finnish Graduate School in Astronomy and Space Physics). We are indebted to Michael Sulzer for suggesting (and practically writing) significant improvements in the introduction.

Topical Editor K. Kauristie thanks M. Lassas and M. Sulzer for their help in evaluating this paper.

## References

Damtie, B., Nygrén, T., Lehtinen, M. S., and Huuskonen, A.: High resolution observations of sporadic-E layers within the polar cap ionosphere using a new incoherent scatter radar experiment,

Ann. Geophys., 20, 1429–1438, 2002,

<http://www.ann-geophys.net/20/1429/2002/>.

Kero, A., Vierinen, J., Enell, C.-F., Virtanen, I., and Turunen, E.: New incoherent scatter diagnostic methods for the heated D-region ionosphere, Ann. Geophys., in press, 2008.

Lehtinen, M., Damtie, B., and Nygrén, T.: Optimal binary phase codes and sidelobe-free decoding filters with application to incoherent scatter radar, Ann. Geophys., 22, 1623–1632, 2004, <http://www.ann-geophys.net/22/1623/2004/>.

Lehtinen, M. S.: Statistical theory of incoherent scatter radar measurements, Ph.D. thesis, University of Helsinki, 1986.

Lehtinen, M. S. and Häggström, I.: A new modulation principle for incoherent scatter measurements, Radio Sci., 22, 625–634, 1987.

Lehtinen, M. S. and Huuskonen, A.: General incoherent scatter analysis and GUISDAP, J. Atmos. Terr. Phys., 58, 435–452, 1996.

Markkanen, M., Vierinen, J., and Markkanen, J.: Polyphase alternating codes, Ann. Geophys., 26, 2237–2243, 2008, <http://www.ann-geophys.net/26/2237/2008/>.

Sulzer, M. P.: A new type of alternating code for incoherent scatter measurements, Radio Sci., 28, 995–1002, 1993.

Vierinen, J., Lehtinen, M., Orispää, M., and Virtanen, I.: Transmission code optimization method for incoherent scatter radar, Ann. Geophys., accepted, 2008a.

Vierinen, J., Lehtinen, M., and Virtanen, I.: Amplitude domain analysis of strong range and Doppler spread radar echos, Ann. Geophys., in press, 2008b.

Virtanen, I. I., Lehtinen, M. S., Nygrén, T., Orispää, M., and Vierinen, J.: Lag profile inversion method for EISCAT data analysis, Ann. Geophys., 26, 571–581, 2007, <http://www.ann-geophys.net/26/571/2007/>.

Virtanen, I., Lehtinen, M., and Vierinen, J.: Towards multi-purpose IS radar experiments, Ann. Geophys., 26, 2281–2289, 2008, <http://www.ann-geophys.net/26/2281/2008/>.