

# Maximum Likelihood Signal Parameter Estimation via Track Before Detect

Murat Üney and Bernard Mulgrew

Institute for Digital Communications, School of Engineering,  
The University of Edinburgh,  
EH9 3JL, Edinburgh, UK  
Emails: {*M.Uney, B.Mulgrew*}@ed.ac.uk

Daniel Clark

School of Eng.& Physical Sciences,  
Heriot-Watt University  
EH14 4AS, Edinburgh, UK  
Email: *D.E.Clark@hw.ac.uk*

**Abstract**—In this work, we consider the front-end processing for an active sensor. We are interested in estimating signal amplitude and noise power based on the outputs from filters that match transmitted waveforms at different ranges and bearing angles. These parameters identify the distributions in, for example, likelihood ratio tests used by detection algorithms and characterise the probability of detection and false alarm rates. Because they are observed through measurements induced by a (hidden) target process, the associated parameter likelihood has a time recursive structure which involves estimation of the target state based on the filter outputs. We use a track-before-detect scheme for maintaining a Bernoulli target model and updating the parameter likelihood. We use a maximum likelihood strategy and demonstrate the efficacy of the proposed approach with an example.

## I. INTRODUCTION

Active sensors send energy packets towards a surveillance region in order to locate objects within from the reflections. For example, radars transmit radio frequency (RF) electromagnetic (EM) pulses and locate reflectors by searching for the pulse waveform in the spatio-temporal energy content of the received signal. This search is often performed using the matched filtering technique in which the received signal is projected onto versions of the transmitted waveform shifted so as to encode the desired reflector locations [1]. The more the energy a projection has, the more likely that it is due to the presence of a reflector.

In this work, we are interested in estimating parameters related to the signal at the output of the matched filters (MF) of a radar. This signal is composed of a distorted version of the waveform auto-correlation function in the presence of a reflector and additive thermal noise. Detection algorithms aim to decide on the existence of an object based on these outputs sampled at selected time instances so as to give the energy of the aforementioned projected signal [2]. These decisions are characterised by a probability of detection and a probability of false alarm which can be found given the energy of the reflected pulse at the receiver front-end  $E$ , and, the noise power  $\beta^2$  (or, the standard deviation of the noise process). These parameters also determine the signal-to-noise ratio which can be used to characterise the expected accuracy in further levels of processing [1].

Often only the noise power  $\beta^2$  is estimated using spatial windows over one snapshot of the outputs from MFs asso-

ciated with different range-bearing bins. These estimates are then used as design parameters in constant-false-alarm rate detection algorithms [3]. However, this approach is prone to errors due to incorrect identification of bins that contain only noise or signal-and-noise as a result of that this identification task requires the temporal information in the received signal which is ignored. Estimation of the signal energy  $E$  requires collection of temporal samples, as well. One way of doing this is to replace the MF bank in the basedband processing chain with iterative processing algorithms preferably working with high sampling rates (e.g., [4], [5]). This places requirements on the hardware architecture that are hard to satisfy in practice.

In this work, we use multiple snapshots from the MF bank collected across a time interval and perform spatio-temporal processing to jointly estimate  $E$  and  $\beta^2$ . We treat the problem as a parameter estimation problem in state space models. This allows us to integrate all the information in the measurements in a single likelihood function for both  $E$  and  $\beta^2$ . Such likelihood functions require the state distribution of the underlying (target) process given the measurement history, which can be found by the prediction stage of Bayesian filtering (or, tracking) recursions [6]. As the measurements are the MF outputs (as opposed to the outputs from a detection algorithm as in widely studied tracking scenarios), the corresponding recursions describe a track-before-detect algorithm.

In this framework, we derive explicit formulae for the parameter likelihood and its score function, i.e., the log-likelihood gradient. We use a maximum likelihood (ML) approach to design an unbiased and minimum variance estimator. In particular, we maximise the log-likelihood by using a coordinate ascent algorithm [7] in which we select the directions of increase based on the gradient and perform (golden section) line search along these directions.

This article is outlined as follows: We give the ML problem definition in Sec. II and detail the parameter likelihood in Sec. III. In Sec. IV, we derive the gradient of the objective function and detail an iterative optimisation procedure. We demonstrate the proposed approach with an example in Sec. V.

## II. PROBLEM DEFINITION

We consider a pulse transmitted towards a surveillance region which gets reflected if it interacts with an object at state

$x = [x_l^T, x_v^T]^T$  where  $x_l$  is the location,  $x_v = \dot{x}_l$  is the velocity of the object and  $(\cdot)^T$  denotes the transpose of a vector. These reflections are sought in the spatio-temporal energy content of the received signal by matched filtering [1]. Typically, the filter output is sampled with a period of the pulse length so as to compute the correlation of the transmitted waveform with the received signal corresponding to the  $i^{\text{th}}$  range bin of width  $\Delta r$  and  $j^{\text{th}}$  bearing bin of width  $\Delta\phi$ . Therefore, at time step  $k$ , the filter output for the bin  $(i, j)$  is given by

$$r_k(i, j) = \langle \mathbf{m}(i, j), \mathbf{r} \rangle \quad (1)$$

where  $\mathbf{r}$  is the (complex) received signal and  $\mathbf{m}(i, j)$  represents the transmitted waveform shifted to the  $(i, j)^{\text{th}}$  bin.

Let  $E_t$  represent the energy of the transmitted pulse, i.e.,  $E_t = \mathbf{m}^H \mathbf{m}$  where  $(\cdot)^H$  denotes the Hermitian transpose. In the presence of a reflecting object at state  $x_k$ , the inner product above leads to

$$r_k(i, j) = E_t e^{j\theta_k} h_{i,j}(x_k) + n_k \quad (2)$$

where  $n_k \sim \mathcal{N}(\cdot; 0, \beta^2)$  is circularly symmetric complex Gaussian noise with complex power  $\beta^2$ ,  $h_{i,j}(x_k)$  specifies the ratio of  $E$  that has been reflected from the  $(i, j)^{\text{th}}$  bin with (an unknown) phase  $\theta_k \sim \mathcal{U}(0, 2\pi)$ .

In this work, we assume a sensor resolution such that an object in the surveillance region affects only a single range-bearing bin, i.e.,

$$h_{i,j}(x_k) = H \delta_{C(x_k), (i,j)} \quad (3)$$

where  $\delta$  is the Kronecker's delta function,  $C: \mathcal{X} \rightarrow M \times N$  maps object states to range-bearing bins and  $H$  is the reflection coefficient.

We consider (2) and (3), and, are interested in estimating the received signal energy  $E \triangleq E_t H$ , and, the noise power  $\beta^2$ . These signal parameters determine the signal-to-noise ratio at the matched filter by

$$SNR = 10 \log_{10} \frac{E^2}{\beta^2}, \quad (4)$$

and are also required to compute false alarm rates and object detection probabilities of threshold rules [1].

We treat these parameters as (non-random) unknown constants and consider an ML solution. We now specify the arguments of the likelihood function: The reflection phase  $\theta_k$  in (2) models the ambiguity related to the exact position of the reflector within the  $(i, j)^{\text{th}}$  bin which cannot be mitigated. The modulus of (2), i.e.,

$$z_k(i, j) \triangleq |r_k(i, j)|$$

neglects the phase and is a sufficient statistic when testing whether  $r_k(i, j)$  is induced by a reflector at state  $x_k$  or noise alone. Therefore, we treat the intensity map given by  $z_k(i, j)$ s as measurements based on which the ML estimation will be relying upon.

Let us denote with  $\mathbf{z}_k$  the concatenation of  $z_k(i, j)$ s for all range-bearing bins, i.e., the intensity map. We would like to base the parameter likelihood on all measurements covering

time step 1 through  $k$ . Therefore, the ML estimation problem can be formulated as

$$(\hat{E}, \hat{\beta}^2) = \arg \max_{E, \beta^2} \log l(\mathbf{z}_1, \dots, \mathbf{z}_k | E, \beta^2), \quad (5)$$

the computation of which will be described next.

### III. THE SIGNAL PARAMETER LIKELIHOOD

Let us represent with a random set  $X_k$ , the events that there exists a reflector with state  $x_k$ , and, none, in which cases  $X_k = \{x_k\}$  and  $X_k = \emptyset$ , respectively.  $X_k$  is referred to as a Bernoulli random finite set (RFS) [8]. The likelihoods for  $z_k(i, j)$  given the signal parameters to be estimated  $E, \beta^2$  for the cases that  $X_k = \{x_k\}$  and  $X_k = \emptyset$  are well known results in the literature: After the uniformly distributed phase is marginalised-out, the modulus  $z_k(i, j)$  is distributed with a Rician distribution, i.e.,

$$\begin{aligned} l_1(z_k(i, j) | x_k; E, \beta^2) &\triangleq p(z_k(i, j) | X_k = \{x_k\}; E, \beta^2) \\ &= \frac{2z_k(i, j)}{\beta^2} \exp\left(-\frac{z_k(i, j)^2 + E^2}{\beta^2}\right) I_0\left(\frac{2z_k(i, j)E}{\beta^2}\right), \end{aligned} \quad (6)$$

if  $X_k = \{x_k\}$  and  $x_k \in C^{-1}(i, j)$ , where  $I_0$  is the zero order modified Bessel function of the first kind. Otherwise,  $z_k(i, j)$  follows a Rayleigh law given by [1, Chp.6]

$$\begin{aligned} l_0(z_k(i, j) | \beta^2) &\triangleq p(z_k(i, j) | X_k = \emptyset; E, \beta^2) \\ &= \frac{2z_k(i, j)}{\beta^2} \exp\left(-\frac{z_k(i, j)^2}{\beta^2}\right). \end{aligned} \quad (7)$$

Let us define the intensity map  $\mathbf{z}_k$  related quantities:

$$\begin{aligned} \Lambda(\mathbf{z}_k | x_k, E, \beta^2) &\triangleq l_1(z_k(C(x_k)) | x_k; E, \beta^2) \\ &\prod_{(i,j) \in \overline{C(x_k)}} l_0(z_k(i, j) | \beta^2), \end{aligned} \quad (8)$$

where  $\overline{C(x_k)}$  denotes the set of range-bearing bins complementing  $C(x_k)$ , and,  $l_0$  and  $l_1$  are given in (6) and (7), respectively. Similarly

$$\Lambda(\mathbf{z}_k | \beta^2) \triangleq \prod_{i,j} l_0(z_k(i, j) | \beta^2) \quad (9)$$

where the product is over all the range-bearing bins.

We assume that the noise processes for different bins and time steps are independent given the state of the object process  $X_k$ . Hence,

$$p(\mathbf{z}_k | X_k, E, \beta^2) = \begin{cases} \Lambda(\mathbf{z}_k | x_k, E, \beta^2), & \text{if } X_k = \{x_k\} \\ \Lambda(\mathbf{z}_k | \beta^2), & \text{if } X_k = \emptyset. \end{cases} \quad (10)$$

We would like to compute the parameter likelihood based on all measurements covering time step 1 through  $k$  which can be found as

$$\begin{aligned} l(\mathbf{z}_1, \dots, \mathbf{z}_k | E, \beta^2) &= \prod_{t=1}^k p(\mathbf{z}_t | \mathbf{z}_{1:t-1}, E, \beta^2) \\ &= l(\mathbf{z}_1, \dots, \mathbf{z}_{k-1} | E, \beta^2) \\ &\quad \times p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2) \end{aligned}$$

after using the chain rule of probabilities in the first line. The recursive structure revealed in the second line is typical to parameter estimation problems in state space models [9], [10], in which the update term is found by marginalising-out the underlying process  $X_k$ , i.e.,

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2) = \int p(\mathbf{z}_k | X_k, E, \beta^2) p(X_k | \mathbf{z}_{1:k-1}, E, \beta^2) \delta X_k \quad (11)$$

where the right hand side is a set integral [8, Chp.11] as  $X_k$  is a set random variable. The first term inside the integral is the likelihood at  $k$  given by (10), and, the second term is the prediction of  $X_k$  based on the previous measurements.

The objective function for the ML solution in (5), hence, is given by

$$J(E, \beta^2; \mathbf{z}_{1:k}) = J(E, \beta^2; \mathbf{z}_{1:k-1}) + \log p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2). \quad (12)$$

Let us now consider the likelihood update term in (11). The computation of the predictive term inside the integral is detailed later in Section III-A. For our discussion on the likelihood update term, suppose that it is given by

$$p(X_k | \mathbf{z}_{1:k-1}, E, \beta^2) = \begin{cases} r_{k|k-1} s_{k|k-1}(x_k), & \text{if } X_k = \{x_k\} \\ 1 - r_{k|k-1}, & \text{if } X_k = \emptyset. \end{cases} \quad (13)$$

Then, (11) expands using the set integration rule in [8, Chp.11] with a Bernoulli RFS characterised by (13) and the likelihood in (10) as

$$\begin{aligned} p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2) &= p(\mathbf{z}_k | \emptyset, E, \beta^2) p(\emptyset | \mathbf{z}_{1:k-1}, E, \beta^2) \\ &+ \int p(\mathbf{z}_k | X_k = \{x_k\}, E, \beta^2) p(X_k = \{x_k\} | \mathbf{z}_{1:k-1}, E, \beta^2) dx_k \\ &= (1 - r_{k|k-1}) \Lambda(\mathbf{z}_k | \beta^2) \\ &+ r_{k|k-1} \int \Lambda(\mathbf{z}_k | x_k, E, \beta^2) s_{k|k-1}(x_k) dx_k. \end{aligned} \quad (14)$$

After dividing both parts of the equation above by  $\Lambda(\mathbf{z}_k | \beta^2)$ , it can easily be shown that

$$\begin{aligned} \log p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2) &= \log \Lambda(\mathbf{z}_k | \beta^2) + \\ &\log(1 - r_{k|k-1} + \\ &r_{k|k-1} \int \frac{l_1(z_k(C(x_k)) | x_k; E, \beta^2)}{l_0(z_k(C(x_k)) | \beta^2)} s_{k|k-1}(x_k) dx_k) \end{aligned} \quad (15)$$

Therefore, the ML objective can be recursively evaluated using (15) in (12). Next, we discuss the computation of the predictive state distribution in (13) using Bernoulli track before detect.

#### A. Bernoulli track before detect

Recursive updating of a Bernoulli object model using measurements with likelihood models in the form of (10) can be carried out using the Bayesian recursive filtering principles [11], [12]. An important component of filtering with RFS models is the object appearance, or, birth model. The probability of object appearance at time  $k$  is given by  $P_b$  and the state of this object is distributed as  $b(x)$ . The probability of

an object that existed at  $k-1$  continuing to exist at time  $k$  is given by  $P_S$ . A Bernoulli model at time  $k-1$  characterised by  $(r_{k-1}, s_{k-1}(x_{k-1}))$ , then, leads to the following prediction:

$$\begin{aligned} r_{k|k-1} &= P_b(1 - r_{k-1}) + r_{k-1} P_S \\ s_{k|k-1}(x) &= \frac{P_b(1 - r_{k-1})}{r_{k|k-1}} b(x) \\ &+ \frac{r_{k-1} P_S}{r_{k|k-1}} \int \pi_{k|k-1}(x | x_{k-1}) s_{k-1}(x_{k-1}) dx_{k-1} \end{aligned} \quad (16)$$

where  $\pi_{k|k-1}$  is the state transition density [11].

Upon receiving  $\mathbf{z}_k$ , the posterior model is given by [11]

$$\begin{aligned} r_k &= \frac{r_{k|k-1} \int g_k(x_k | E, \beta^2) dx_k}{1 - r_{k|k-1} + r_{k|k-1} \int g_k(x_k | E, \beta^2) dx_k} \\ s_k(x_k) &= \frac{g_k(x_k | E, \beta^2)}{\int g_k(x_k | E, \beta^2) dx_k} \\ g_k(x_k | E, \beta^2) &\triangleq \frac{l_1(z_k(C(x_k)) | x_k; E, \beta^2)}{l_0(z_k(C(x_k)) | \beta^2)} s_{k|k-1}(x_k). \end{aligned} \quad (17)$$

As a result, the predictive term in (13) required for the parameter likelihood is found by iterating prediction and update cycles using (16) and (17).

The object birth model we use is selected so as to have a uniform distribution in the location component which is nonzero in the sensor field of view, and, a uniform distribution in the velocity component which is nonzero if the speed is between selected minimum and maximum values:

$$b([x_l^T, x_v^T]^T) = \mathcal{U}_{FOV}(x_l) \mathcal{U}_{v_{min} \leq |x_v| \leq v_{max}}(x_v). \quad (18)$$

#### IV. MAXIMUM LIKELIHOOD SIGNAL AMPLITUDE AND NOISE POWER ESTIMATION

The objective function of the ML solution in (5) is not straightforward to maximise partly because, in practice, only a noisy approximation of it can be obtained via particle methods, and, the term due to  $E$  in the right hand side (RHS) of (15) is dominated by the influence of  $\beta^2$  through the first term (see, e.g., the example in Section V).

Our ML realisation strategy is to apply coordinate ascent along the projections of the log-likelihood gradient. This can be viewed as a subgradient approach with the difference that we do not explicitly select step sizes to move along the selected subgradient direction. Instead, we perform (golden ratio) line search along the subgradient direction. The gradient equals the sum of the gradients of the log update term in (15) over time. Note that, the first term in the RHS of (15) is independent of  $E$ , hence,

$$\begin{aligned} \frac{\partial \log p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2)}{\partial E} &= \frac{r_{k|k-1} \int \frac{\partial \log l_1}{\partial E} \frac{l_1}{l_0} s_{k|k-1}}{1 - r_{k|k-1} + r_{k|k-1} \int \frac{l_1}{l_0} s_{k|k-1}} \\ &= \frac{r_{k|k-1} \int \frac{\partial \log l_1}{\partial E} g_k}{1 - r_{k|k-1} + r_{k|k-1} \int g_k} \end{aligned} \quad (19)$$

where the first line follows after differentiating the term inside the integral and using the identity  $\partial l_1 / \partial E = l_1 \times \partial \log l_1 / \partial E$ . The second line is obtained through the definition of  $g_k$  in (17).

The partial derivative of  $\log l_1$  with respect to  $E$  can easily be found [13] as

$$\frac{\partial \log l_1}{\partial E} = -\frac{2E}{\beta^2} + \frac{2z_k(C(x_k))}{\beta^2} \frac{I_1(2z_k(C(x_k))E/\beta^2)}{I_0(2z_k(C(x_k))E/\beta^2)} \quad (20)$$

where  $I_0$  and  $I_1$  are the modified Bessel functions of the first kind of order zero and one, respectively. After substituting (20) in (19), we get

$$\frac{\partial \log p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2)}{\partial E} = -\frac{2Er_k}{\beta^2} + \frac{2r_{k|k-1} \int \frac{z_k I_1}{\beta^2 I_0} g_k}{1 - r_{k|k-1} + r_{k|k-1} \int g_k}, \quad (21)$$

where  $r_k$  and  $g_k$  are given in (17).

Next, we consider the partial derivative of the log update term with respect to  $\beta^2$ :

$$\frac{\partial \log p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2)}{\partial \beta^2} = \sum_{i,j} \frac{\partial \log l_0}{\partial \beta^2} + \frac{r_{k|k-1} \int \left( \frac{\partial \log l_1}{\partial \beta^2} - \int \frac{\partial \log l_0}{\partial \beta^2} \right) g_k}{1 - r_{k|k-1} + r_{k|k-1} \int g_k} \quad (22)$$

The first partial derivative inside the integral above can be found [13] as

$$\frac{\partial \log l_1}{\partial \beta^2} = -\frac{1}{\beta^2} \left( 1 - \frac{z_k^2 + E^2}{\beta^2} + \frac{2z_k E}{\beta^2} \frac{I_1(2z_k E/\beta^2)}{I_0(2z_k E/\beta^2)} \right),$$

and the derivative of log-noise term with respect to the noise power is given by

$$\frac{\partial \log l_0}{\partial \beta^2} = -\frac{1}{\beta^2} + \frac{z_k^2}{\beta^4}.$$

Therefore, the first term inside the integral in (22) is given by

$$\frac{\partial \log l_1}{\partial \beta^2} - \frac{\partial \log l_0}{\partial \beta^2} = \frac{E^2}{\beta^4} - \frac{2Ez_k}{\beta^4} \frac{I_1}{I_0}. \quad (23)$$

As a result, the gradient of the log-likelihood function in (12) can be computed using (19)–(23) in the recursive form given by

$$\nabla J(E, \beta^2; \mathbf{z}_{1:k}) = \nabla J(E, \beta^2; \mathbf{z}_{1:k-1}) + \left[ \frac{\partial \log p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2)}{\partial E} \frac{\partial E}{\partial \beta^2} + \frac{\partial \log p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, E, \beta^2)}{\partial \beta^2} \right] \quad (24)$$

In order to maximise  $J$ , we adopt a coordinate ascent approach: Starting from an initial point  $(E_0, \beta_0^2)$ , at each iteration  $m$ , we find the gradient vector (24) and select a line search direction as follows:

$$d_m = \frac{\nabla J(E, \beta^2; \mathbf{z}_{1:k})^T e_m}{\|\nabla J(E, \beta^2; \mathbf{z}_{1:k})^T e_m\|} \quad (25)$$

$$e_m = \begin{cases} [0, 1]^T, & \text{if } m = 0, 2, 4, \dots \\ [1, 0]^T, & \text{if } m = 1, 3, 4, \dots \end{cases}$$

Then, we solve the following one dimensional problem

$$(E_{m+1}, \beta_{m+1}^2) = \arg \min_{\lambda \in [0, \lambda_{max}]} J([E_m, \beta_m^2]^T + \lambda d_m; \mathbf{z}_{1:k}) \quad (26)$$

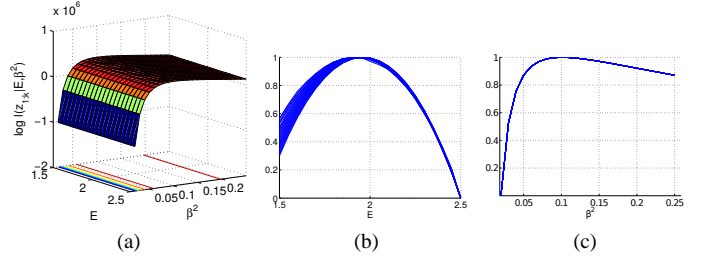


Fig. 1. (a) Log-likelihood surface for the example scenario evaluated over the grid  $1.5 \leq E \leq 2.5$  and  $0.01 \leq \beta^2 \leq 0.25$  with step sizes of 0.05 and 0.01, respectively. (b) The normalised profile of the surface along  $E$ -axis for varying  $\beta^2$ . (c) The normalised profile along  $\beta^2$  for varying  $E$ .

using golden section search [14].

The iterations terminate if two consecutive points are closer than a selected tolerance value. In other words, the solution to (5) is declared as  $(\hat{E}, \hat{\beta}^2) = (E_{m+1}, \beta_{m+1}^2)$  if

$$\|[E_{m+1}, \beta_{m+1}^2]^T - [E_m, \beta_m^2]^T\| < \delta.$$

The computations are carried out using particle methods: We sample from the predictive distribution in (13) using a Sequential Monte Carlo realisation [11] of the recursions given by (16) and (17). In order to sample from the birth model in (18), we find a grid of  $L$  samples for each bin  $C^{-1}(i, j)$  and concatenate with velocity components generated from  $\mathcal{U}_{v_{min} \leq |x_v| \leq v_{max}}$ . The integrals in (19)–(23) are estimated using samples generated from  $g_k$  during Bernoulli track before detect within the Monte Carlo method [15, Chp.3].

## V. EXAMPLE

Let us consider an example scenario consisting of an object with initial state  $x_0 = [-503.5, 4974.6, 5, 0]^T$  moving in accordance with a constant velocity motion model with a small process noise term and a sensor located at the origin with range and bearing resolutions of 10m and  $1^\circ$ , respectively. The sensor field of view is selected as the region bounded by 4800m and 5200m in range and  $\pm 10^\circ$  around the  $y$ -axis in order to restrict the size of the intensity map and hence the volume of computations needed. As a result, a  $40 \times 20$  intensity map is observed for  $k = 200$  time steps with signal amplitude  $E = 2$  and noise power  $\beta^2 = 0.1$ . The target signal, hence, has  $\sim 16$ dB SNR in the corresponding bin.

In Fig. 1(a), we present the ML objective surface evaluated for a given realisation of the measurement history  $\mathbf{z}_{1:200}$  and a grid of  $(E, \beta^2)$  values using the parameter likelihood detailed in Section III together with MC computations. For Bernoulli track before detect, we use  $L = 225$  uniform grid points for each of the 800 state bins given by  $C^{-1}(i, j)$  to represent  $b(x)$  in (16). Here,  $\pi$  is selected as a constant velocity motion model with a small additive process noise term. We maintain 10000 samples generated from  $s_k$  in (17) after resampling weighted samples from  $g_k$ . The probability of birth  $P_b = 0.01$  whereas the  $P_S = 0.99$ .

There are much less target associated measurements contributing to (15) compared to noise associated terms. This manifests itself in Fig. 1(a) as the significantly smaller curvature of the surface along  $E$  compared to that along  $\beta^2$ .

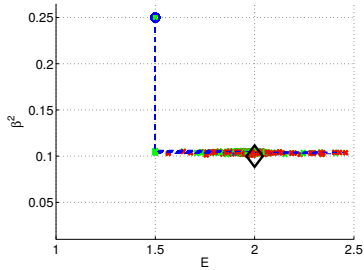


Fig. 2. Iterative maximisation of the parameter likelihood for 200 Monte Carlo runs: Initial point (blue circle), and converged estimates (red crosses) joined by estimates in the intermediate steps (green crosses and dashed lines). Diamond is the correct value of  $(E, \beta^2)$ .

Correspondingly, the Cramer-Rao lower bound (CRLB) for  $E$  is higher and estimates of  $E$  will be less accurate [16, App. 8A]. For a closer look to the log-likelihood objective function, we provide the normalised profiles of the surface along  $E$  and  $\beta^2$  axis in Fig. 1(b) and (c), respectively. Note that, for almost all values of  $E$  in the grid, it is possible to estimate  $\beta^2$  through maximising the objective along  $\beta^2$  direction.

Next, we use the proposed scheme described in Section IV starting from  $(E_0, \beta_0^2) = (1.5, 0.25)$ . After evaluating the gradient in (24), a golden section line search is performed along the direction selected using (25). The initial search interval length, i.e.,  $\lambda_{max}$  in (26), is selected to ensure that the resulting interval of uncertainty will be smaller than 0.001. For golden ratio search, this is given by  $\lceil \log(0.001 - 0.618) \rceil + 1$ . The tolerance  $\delta$  for checking convergence is selected as 0.001.

We repeat the proposed ML scheme for 200 Monte Carlo realisations of this scenario. The iterations are depicted in Fig. 2, in which the resulting estimates are shown by red crosses. The average number of iterations is 3.6. The average of  $(\hat{E}, \hat{\beta}^2)$  is  $(1.972, 0.104)$ , which is very close to the true value of  $(2, 0.1)$ . This indicates that the optimisation scheme is approximately unbiased. The variance of the estimation scheme as measured by the empirical average of the squared errors for  $\hat{E}$  and  $\hat{\beta}^2$  are found as  $1.78 \times 10^{-2}$  and  $1.7 \times 10^{-5}$ , respectively. The variance of  $\hat{E}$  is much higher than that for  $\hat{\beta}^2$ , as expected.

## VI. CONCLUSIONS AND FUTURE WORK

In this work, we proposed a ML scheme for jointly estimating the signal amplitude and noise power using the matched filter outputs of an active sensor. This likelihood involves track-before-detect in order predict the underlying target process which is modelled by a Bernoulli RFS. We derived explicit formulae for the score function and proposed an iterative maximisation procedure using a coordinate ascent approach.

It is possible to improve the optimisation step of the proposed algorithm. For example, Newtonian methods can be used after finding the Hessian of the log-likelihood. Such an approach would remove the need for line search and potentially provide more accurate estimates for fewer number of iterations and objective evaluations.

There is a tradeoff between the observation length and the accuracy in which  $E$  can be estimated (see, e.g., [17]). The relations between the Hessian, Fisher information and the associated Cramer-Rao lower bound can be explored in order to investigate this tradeoff. Another possible extension of this work is to accommodate multi-Bernoulli models in order to handle multiple moving objects via track-before-detect [18]. Estimation of parameters for other Swerling target types [1] within the proposed ML framework is also left as future work.

## ACKNOWLEDGMENT

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) grants EP/J015180/1 and EP/K014277/1, and the MOD University Defence Research Collaboration in Signal Processing.

## REFERENCES

- [1] M. A. Richards, *Fundamentals of Radar Signal Processing*. McGraw Hill, 2005.
- [2] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall, 1998, vol. II.
- [3] H. Rohling, "Radar cfar thresholding in clutter and multiple target situations," *IEEE Trans. on Aerospace and Elec. Sys.*, vol. AES-19, no. 4, pp. 608–621, July 1983.
- [4] R. C. Raup, R. A. Ford, G. R. Krumpolz, M. G. Czerwinski, and T. E. Clark, "The best approximation of radar signal amplitude and delay," *Lincoln Laboratory Journal*, vol. 3, pp. 311–327, 1990.
- [5] A. Sheikhi, M. Nayebi, and M. Aref, "Adaptive detection algorithm for radar signals in autoregressive interference," *IEE Proceedings - Radar, Sonar and Nav.*, vol. 145, no. 5, pp. 309–314, Oct 1998.
- [6] O. Cappé, S. J. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proceedings of the IEEE*, vol. 95, pp. 899–924, 2007.
- [7] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
- [8] R. P. S. Mahler, *Statistical Multisource Multitarget Information Fusion*. Springer, 2007.
- [9] M. Uney, B. Mulgrew, and D. Clark, "Cooperative sensor localisation in distributed fusion networks by exploiting non-cooperative targets," in *IEEE Stat. Sig. Proc. Workshop (SSP)*, 2014, June 2014, pp. 516–519.
- [10] —, "Target aided online sensor localisation in bearing only clusters," in *Sensor Sig. Proc. for Defence (SSPD)*, 2014, Sept 2014, pp. 1–5.
- [11] B. Ristic, B.-T. Vo, B.-N. Vo, and A. Farina, "A tutorial on bernoulli filters: Theory, implementation and applications," *Signal Processing, IEEE Transactions on*, vol. 61, no. 13, pp. 3406–3430, July 2013.
- [12] M. Rutten, N. Gordon, and S. Maskell, "Recursive track-before-detect with target amplitude fluctuations," *Radar, Sonar and Navigation, IEE Proceedings -*, vol. 152, no. 5, pp. 345–352, October 2005.
- [13] C. Carobbi and M. Cati, "The absolute maximum of the likelihood function of the rice distribution: Existence and uniqueness," *IEEE Trans. on Instrumentation and Meas.*, vol. 57, no. 4, pp. 682–689, April 2008.
- [14] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear programming: Theory and algorithms*, 3rd ed. John Wiley & Sons, 2006.
- [15] G. Casella and C. P. Robert, *Monte Carlo Statistical Methods*, 2nd ed. Springer, 2005.
- [16] S. Kay, *Fundamentals of Statistical Signal Processing, Volume III: Practical Algorithm Development*. Pearson Education, 2013.
- [17] D. Clark, B. Ristic, B.-N. Vo, and B. T. Vo, "Bayesian multi-object filtering with amplitude feature likelihood for unknown object snr," *IEEE Trans. on Signal Proc.*, vol. 58, no. 1, pp. 26–37, Jan 2010.
- [18] B.-N. Vo, B.-T. Vo, N.-T. Pham, and D. Suter, "Joint detection and estimation of multiple objects from image observations," *IEEE Trans. on Signal Processing*, vol. 58, no. 10, pp. 5129–5141, Oct 2010.