23rd International Symposium on Mathematical Theory of Networks and Systems
Hong Kong University of Science and Technology, Hong Kong, July 16-20, 2018

On the Uniqueness of Maximum Likelihood Estimation of Nuclear Sources

J. Haug, M. F. Anjum, B. D. O. Anderson, S. Dasgupta, R. Mudumbai and H. E. Baidoo-Williams

*Abstract***— In an earlier paper, [1], we have considered the Maximum Likelihood (ML) localization of a stationary nuclear source using the time of arrival of particles modeled as a Poisson process. In this paper we consider whether the ML location estimate characterized in [1] is unique. In particular, we show that the question of uniqueness is equivalent to whether or not the root locus of a certain transfer function admits a single pair of nonzero imaginary axis crossing.**

*Index Terms***— Nuclear source, maximum likelihood localization, Poisson, root locus.**

I. Introduction

Identification and tracking of nuclear materials has recently attracted significant research [1]- [5], and is challenging as nuclear materials vary widely in concentrations and compositions [5], are concealed by shielding material, and submerged in background radiation [6]. Detectors [7] and sensors used in practical tracking systems, also vary widely in their quality. Ultimately the detection of nuclear radiation comprises a sequence of events, modeled as a Poisson arrival process [8], [9], involving the absorption of discrete particles.

Source localization with Poisson models has been studied by [1] and [2] among others. While [2] uses a non-concave expectation maximization process, [1] uses Maximum Likelihood (ML) localization. In particular [1] considers an ideal detector, free from timing inaccuracies caused by the quantum energy-time uncertainty principle, [11], that travels with a known fixed velocity. It formulates a likelihood function for localizing a stationary source from a finite set of arrival times, $t_1, \cdots, t_n$, measured over an infinite horizon.

A key question unanswered in [1] is whether the likelihood function is unimodal. One should note that for $n > 2$, simulations indicate that the likelihood function is indeed unimodal. In this paper we make strides towards resolving this question by relating unimodality to a root locus problem. This permits the use of standard root locus methods to pinpoint the maximum. Further, rules of thumb for root loci suggest the uniqueness of this maximum.

Haug, Anjum, Mudumbai and Dasgupta are with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242, USA. Dasgupta is also with the Shandong Computer Science Center, Shandong Provincial Key Laboratory of Computer Networks, China. Anderson is with the Research School of Engineering, Australian National University, with Hangzhou Dianzi University, Hangzhou, China and with Data61-CSIRO in Canberra, A.C.T., Australia. Baidoo-Williams is with Vocal Technologies, 520 Lee Entrance, Suite 202 Amherst, New York 14228. E:mail: [joshua-haug,md-fahim]@uiowa.edu, [rmudumbai,dasgupta]@engineering.uiowa.edu, Brian.Anderson@anu.edu.au and hbaidoow@gmail.com,. This work was supported in part by NSF grants ECCS-1150801, CNS-1329657 and CCF-1302456 and ONR grant N00014-13-1-0202.

Section II provides the log likelihood function derived in [1]. Section III gives a reinterpretation of the maximum which has an interesting geometrical implication. Section IV establishes the equivalence with a root locus problem.

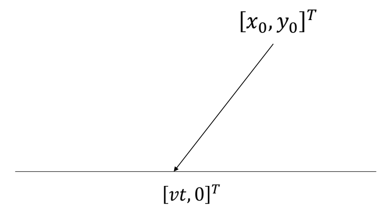

Fig. 1. Ideal detector moving in a straight line with constant velocity.

II. Problem formulation

Suppose the source is at $[x_0, y_0]^T \in \mathbb{R}^2$ and the detector travels at a constant velocity along the x axis, i.e. with a known velocity $[v, 0]^T$, see Figure 1. The time of arrival of particles at the detector is an inhomogeneous Poisson process, [15], with mean arrival rate:

$$\lambda(t) = \frac{A_0}{y_0^2 + (x_0 - vt)^2}, \tag{1}$$

where A_0 is a source strength parameter that depends on the type, shape and volume of the nuclear source as well as detector characteristics. This precludes directional detectors like CZT Compton scattering devices [13] but is rather consistent with scintillation devices [14] in isotropic media.

Assuming that a set of independent arrival times $\{t_1, \cdots, t_n\}$ observed over an interval $[T_1, T_2]$, [1] shows that the conditional density $f(t_1, \cdots, t_n | A_0, x_0, y_0)$ obeys

$$\arg\max_{A_0, x_0, y_0} \log(f(t_1, \cdots, t_n | A_0, x_0, y_0)) = \arg\max_{A_0, x_0, y_0} L(A_0, x_0, y_0)$$

where the modified log likelihood function obeys

$$L(A_0, x_0, y_0) = -\int_{T_1}^{T_2} \lambda(t) dt + \sum_{i=1}^{n} \log \lambda(t_i).$$

This assumes an ideal detector detecting particle counts over vanishingly small intervals. Practical sensors with detection time bins as short as 150 ms exist, [12], and approach the performance of an ideal sensor. As for large t, $\lambda(t)$ in (1) approaches zero, this ensures a finite n. We study an observation interval (T_1, T_2) that extends over the entire real axis i.e. with

$$\lim_{T_1\to-\infty}\lim_{T_2\to\infty} L(A_0,x_0,y_0) = \frac{\pi A_0}{vy_0} + n\log A_0 + \sum_{i=1}^{n}\log\left(\frac{1}{y_0^2+(x_0-vt_i)^2}\right).$$

Thus, A, x and y, defined as the ML estimates of A_0, x_0 and y_0, must respectively satisfy the critical equations

$$A = \frac{nvy}{\pi}, \tag{2}$$

$$\sum_{i=1}^{n}\frac{x-vt_i}{y^2+(x-vt_i)^2} = 0, \tag{3}$$

$$\sum_{i=1}^{n}\frac{1}{y^2+(x-vt_i)^2} = \frac{n}{2y^2}. \tag{4}$$

From (2) we see that the uniqueness of these estimates hinges on (3) and (4) having a unique solution. As neither these nor the physical set up distinguish between $\pm y$ and $y=0$ is meaningless, uniqueness requires that a (x,y), $y>0$ satisfy (3) and (4).

This is satisfied if there is a *unique* $x\in\mathbb{R}$ for which there is a *positive real* y such that (3) and (4) are satisfied; and if at this unique x the $y>0$ satisfying (3) and (4) is unique. Also note that for $n=2$, there is no unique solution as any $y^2 = |x-vt_1||x-vt_2|$ satisfies both (3) and (4). This is unsurprising as three quantities cannot be estimated from two observations. Thus we will focus on $n>2$.

III. Reformulation

We now provide two reformulations of the simultaneous equations (3) and (4). The first provides a geometric interpretation that we find attractive. The second leads in Section IV to a root locus view point.

Observe that (3) and (4) simultaneously hold for $y\neq 0$ iff

$$\begin{aligned} & \textstyle\sum_{i=1}^{n}\frac{x-vt_i}{y^2+(x-vt_i)^2} = 0 \text{ and } \sum_{i=1}^{n}\frac{y}{y^2+(x-vt_i)^2} = \frac{n}{2y} \\ \Leftrightarrow\quad & \textstyle\sum_{i=1}^{n}\frac{x-vt_i-jy}{y^2+(x-vt_i)^2} - \frac{n}{2jy} = 0 \\ \Leftrightarrow\quad & \textstyle\frac{1}{2jy}\sum_{i=1}^{n}\frac{x-vt_i-jy}{x-vt_i+jy} = 0 \Leftrightarrow \sum_{i=1}^{n}\frac{jy-x+vt_i}{x-vt_i+jy} = 0, \end{aligned} \tag{5}$$

where the last equality uses the fact that (x,y) solves (3, 4) iff so does $(x,-y)$. As each summand in (5) has magnitude 1, with $\theta_i = \arctan[y/(x-vt_i)]$, (5) is equivalent to

$$\sum_{i=1}^{n} e^{j2\theta_i} = 0. \tag{6}$$

As depicted in Figure 2 for $n=3$, the θ_i represent the angle made by the source with the x-axis at each point of detection. That (6) holds at these angles is worthy of future exploration.

We next set up the promised root locus framework.

Theorem 3.1: For a given set of times $t_1<t_2<\cdots<t_n$, and a given real s define $c_i = vt_i$ and $c=[c_1,\cdots,c_n]^T$. Define the polynomial

$$p(s,c,K) = \sum_{i=1}^{n}(s+K-c_i)\prod_{k\neq i}(s+c_k-K),\ K\in\mathbb{R}. \tag{7}$$

Then real x and $y>0$ solve (3) and (4) iff $p(jy,c,x)=0$.

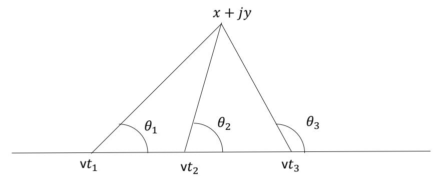

Fig. 2. Depiction of θ_i when $n=3$.

Proof: As the foregoing shows, any (x,y), $y\neq 0$, that solve (3) and (4) obey (5). As $y\neq 0$, this is equivalent to $p(jy,c,x)=0$ for $y>0$. ∎

Thus (3) and (4) have a unique solution, with real x and $y>0$ iff there is one and only one real $K=K^*$ such that $p(s,c,K^*)$ has nonzero imaginary roots and there is exactly one such pair. If this pair of roots are $\pm jy^*$ then $[x,y]=[K^*,\pm y^*]$ are the only solutions of (3) and (4) for which $y\neq 0$. For ease of calculation below, note that if a $c_i<0$, we could replace each c_i by c_i+L and K by $K+L$ for a sufficiently large L to make the following hold without loss of generality.

Assumption 3.1: For all i, $c_i = st_i > 0$.

IV. A Root Locus View

In view of Theorem 3.1 we focus on the polynomial in (7). We have the following lemma.

Lemma 4.1: With $z=s-K$ and

$$f(s,c,K) = \prod_{i=1}^{n}(s+c_i-K), \tag{8}$$

the polynomial in (7) obeys

$$\begin{aligned} p(s,c,K) &= 2sf'(s,c,K) - nf(s,c,K) & (9)\\ &= 2Kf'(z,c,0) + p(z,c,0) & (10)\end{aligned}$$

Proof: Equation (9) follows from:

$$\begin{aligned} p(s,c,K) &= \sum_{i=1}^{n}(s+K-c_i)\prod_{k=1,k\neq i}^{n}(s-K+c_k) \\ &= \sum_{i=1}^{n}(2s-(s-K+c_i))\prod_{k=1,k\neq i}^{n}(s-K+c_k) \\ &= 2s\sum_{i=1}^{n}\prod_{k=1,k\neq i}^{n}(s-K+c_k) - n\prod_{k=1}^{n}(s-K+c_k) \\ &= 2sf'(s,c,K) - nf(s,c,K) \end{aligned}$$

Now $f(s,c,K) = \prod_{i=1}^{n}(s-K+c_i) = f(s-K,c,0)$ and consequently $f'(s,c,K) = \frac{\partial f(z,c,0)}{\partial z}|_{z=s-K} = f'(s-K,c,0)$. This means that with $z=s-K$, (10) follows from (9) and the following calculation:

$$\begin{aligned} p(s,c,K) &= 2sf'(s,c,K) - nf(s,c,K) \\ &= 2sf'(s-K,c,0) - nf(s-K,c,0) \\ &= 2(z+K)f'(z,c,0) - nf(z,c,0) \\ &= 2Kf'(z,c,0) + 2zf'(z,c,0) - nf(z,c,0) \\ &= 2Kf'(z,c,0) + p(z,c,0). \end{aligned}$$

∎

Lemma 4.1 establishes an equivalence between the problem of showing there is a single (x,y) with $y>0$ for which $p(jy,c,x)=0$, and a standard root locus problem, which we now outline. Define with $z=s-K$,

$$q(z,c,K)=p(s,c,K)$$

Now observe that for (3) to hold, at least one summand must be negative and another positive. Under Assumption 3.1, this must mean that x must be positive. Thus as $K=x$ in Theorem 3.1, (3) can only hold if $K>0$. Further the zeros of $q(z,c,K)$ are identical to those of

$$h(z,c,K)=1+K\frac{2f'(z,c,0)}{p(z,c,0)}. \tag{11}$$

Finding x and y thus reduces to a standard root locus problem for the transfer function $h(z,c,K)$. Specifically, we need to see if with z on this root locus $z+K=s$ crosses the positive imaginary axis at only a single point. We now explore the properties of this root locus. The theorem below shows that this root locus can have at most one pair of complex roots.

Theorem 4.1: Under the notation of Theorem 3.1 suppose $c_1>c_2>\cdots>c_n$, $n>2$ and that Assumption 3.1 holds. Then the polynomial $p(s,c,K)$ has at most one pair of complex zeros which arise only if $c_1-K>0>c_n-K$.

Proof: Define $b_i=c_i-K$ and consider two cases.

Case I: $0\notin\{b_1,\cdots,b_n\}$**.** Observe $-b_i$ are the zeros of $f(s,c,K)$. For any interval $(-b_k,-b_{k+1})$ for which $-b_k,-b_{k+1}$ have the same sign, there will hold $s\neq 0$ throughout the interval.

Without loss of generality (as will be evident from the argument below), suppose the selected adjacent pair of zeros of $f(s,c,K)$ both have positive sign, i.e. $0<-b_k<-b_{k+1}$. Now because all zeros of $f(s,c,K)$ are real, so are all zeros of its derivative, $f'(s,c,K)$, and in fact the zeros of the derivative interlace those of $f(s,c,K)$. This is easily seen by Rolle's theorem.

There is precisely one zero of $f'(s,c,K)$ lying in the interval $(-b_k,-b_{k+1})$. Therefore f' and also, because $s\neq 0$ in the interval $(-b_k,-b_{k+1})$, the polynomial $sf'(s,c,K)$ assume opposite signs at the points $-b_k,-b_{k+1}$. Therefore, in light of (9) we see that the polynomial $p(s,c,K)$ assumes opposite signs at the same points $-b_k,-b_{k+1}$. Therefore the polynomial has at least one zero, and necessarily an odd number of zeros, in the interval $(-b_k,-b_{k+1})$.

If the b_i have different signs, there are precisely $n-2$ intervals $(-b_k,-b_{k+1})$ which do not include the origin, and therefore there are at least $n-2$ real zeros of $p(s,c,K)$, lying in these intervals. Since p has degree n, this means it can have at most 2 complex zeros.

If the b_i have the same sign for all i, there are $n-1$ intervals such that each contains one zero of $p(s,c,K)$. None could contain an odd number greater than one, since there are n zeros in all. Since also all but one zero are guaranteed real, the remaining zero must be real, assuming a value less than $-b_1$ or greater than $-b_n$. Thus for $p(s,c,K)$ to have a complex zero there must be a b_i that is negative and another that is positive.

Case II: $0\in\{b_1,\cdots,b_n\}$**.** Now suppose some $b_i=0$. As the b_i are distinct this is the *only* b_i that is zero. Thus dropping the arguments c and K one can write

$$f(s)=sg(s) \tag{12}$$

where the degree $n-1$ polynomial $g(s)$ has $n-1$ real, distinct and nonzero roots.Further under (12), (9) reduces to

$$\begin{aligned} p(s) &= 2sf'(s)-nf(s) \\ &= 2s^2g'(s)-2sg(s)-nsg(s) \\ &= s\left(2sg'(s)-(n-2)g(s)\right). \end{aligned} \tag{13}$$

As for $n>2$, $n-2>0$, the argument in case I shows that $2sg'(s)-(n-2)g(s)$ has at least $n-3$ real zeros, and thus $p(s)$ has has at least $n-2$ real zeros with one at zero. Observe $g(s)$ inherits all nonzero roots of $f(s)$. Thus as argued in case I for $2sg'(s)-(n-2)g(s)$ and hence $p(s,c,K)$ to have a complex zero there must be a b_i that is negative and another that is positive. ■

The next theorem sets up a relatively easy construction of the root locus.

Theorem 4.2: Consider $f'(z,c,0)$ and $p(z,c,0)$ in (10), under Assumption 3.1, i.e. with $0<c_1<c_2<\cdots<c_n$. Define β_i, ordered such that $\beta_{n-1}<\beta_{n-2}<\cdots<\beta_1$, to be the zeros of $f'(z,c,0)$ and α_i, ordered such that $\alpha_n<\alpha_{n-1}<\cdots<\alpha_1$, to be the zeros of $p(z,c,0)$. Then there holds:

$$\beta_{n-1}<\alpha_n<\beta_{n-2}<\cdots<\beta_1<\alpha_2<0<\alpha_1. \tag{14}$$

Proof: Observe with $\beta_i<0$,

$$f'(z,c,0)=n(z-\beta_1)\cdots(z-\beta_{n-1}).$$

Thus

$$p(z,c,0)=n\left(2z(z-\beta_1)\cdots(z-\beta_{n-1})-(z+c_1)\cdots(z+c_n)\right). \tag{15}$$

Recall from Rolle's theorem that

$$-c_n<\beta_{n-1}<-c_{n-1}<\cdots<\beta_1<-c_1<0. \tag{16}$$

Now in (15) for sufficiently large positive z, $p(z,c,0)>0$. Further $p(0,c,0)<0$. Thus indeed $\alpha_1>0$.

Now call $\beta_0=0$. Observe that the first summand of $p(z,c,0)$ given in (15) is zero for each $z=\beta_i$, for all $i\in\{0,\cdots,n-1\}$. Thus at each β_i

$$p(\beta_i,c,0)=-(\beta_i+c_1)\cdots(\beta_i+c_n).$$

Put differently at each $z=\beta_i$, $p(z,c,0)$ equals $-f(z,c,0)=-(z+c_1)\cdots(z+c_n)$. As the latter has a solitary root between consecutive β_i, (see (16)) it follows that for each i, $p(\beta_i,c,0)p(\beta_{i+1},c,0)<0$, i.e. a sign change occurs in $p(z,c,0)$ between consecutive β_i. Thus there is a zero α_{i+2} of $p(z,c,0)$ in each interval (β_{i+1},β_i). As α_1 is outside these intervals, and $p(z,c,0)$ has degree n, there is precisely one root of $p(z,c,0)$ in each interval (β_{i+1},β_i) and these together with α_1 are the only roots of $p(z,c,0)$. The result follows. ■

The implications of Theorem 4.2 can be gleaned from the well-known rules of thumb for root locus construction, [16].

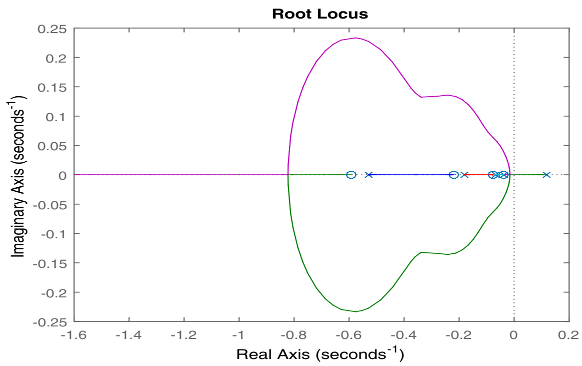

Fig. 3. Root locus of $h(z,c,K)$. The $\times$ are the poles of $h(z,c,K)$ and o the roots of $f'(z,c,0)$.

Poles chase zeros, and all points to the right of an even number of poles and zeros are in the root locus. Thus there are break points in the root locus of $h(z,c,K)$ for $K>0$, in the intervals $(-\infty,\beta_{n-1})$ and (α_2,α_1). This is illustrated in Figure 3 which depicts the root locus of $h(z,b,K)$ for an $n=5$ example.

Conjecture For z on the root locus of $h(z,c,K)$, $z+K$ crosses the positive imaginary axis precisely once.

Regardless of the validity of this conjecture this equivalence between the solutions of (3) and (4) and the root locus of $h(z,c,K)$ provides a means to solving finding the solution of these equations. Specifically, using Matlab's *rlocus* program construct the root locus of $h(z,c,K)$, shift it to the right everywhere by the corresponding K and determine the values of K at which positive imaginary axis is crossed. These estimates so obtained can be refined if needed by a Newton-Raphson algorithm initialized with this estimate. Thus consider with $n=9$, the randomly selected $c=[1.1140,2.1875,2.5294,3.2589,3.6232,3.6535,3.8300,0.3902,0.5079]^T$. With z on the part of the root locus above the real axis the $z+K$ is plotted in Figure 4. Evidently the imaginary axis is crossed only once in the upper half plane. Refined plot reveals $y=.9272$ and $x=.284$.

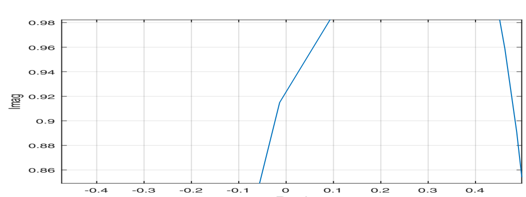

Fig. 4. Portion of the Root locus of $h(z,c,K)$ in the upper half plane, for the given example.

V. Conclusion

In this extended abstract, we have shown an intriguing equivalence between the ML localization of nuclear radiation and a root locus problem. We have argued that this equivalence can be exploited to obtain ML location estimate. Rules of thumb of root locus construction suggest that this estimate is unique.

References

[1] H. E. Baidoo-Williams, R. Mudumbai, E. Bai and S. Dasgupta, "Some theoretical limits on nuclear source localization and tracking", in *Proceedings of Information Theory and Applications Workshop,* 2015.
[2] N. S. V. Rao, M. Shankar, J.-C. Chin, D. Yau, S. Srivathsan, S. Iyengar, Y. Yang, and J. Hou, "Identification of low-level point radiation sources using a sensor network," in *International Conference on Information Processing in Sensor Networks (IPSN)*, 2008.
[3] R. Nemzek, J. Dreicer, D. Torney, and T. Warnock, "Distributed sensor networks for detection of mobile radioactive sources," *Nuclear Science, IEEE Transactions on*, pp. 1693–1700, Aug 2004.
[4] H. Baidoo-Williams, S. Dasgupta, R. Mudumbai, and E. Bai, "On the gradient descent localization of radioactive sources," *IEEE Signal Processing Letters,* pp. 1046 - 1049, 2013.
[5] P. Kump, E.-W. Bai, K.-S. Chan, and W. Eichinger, "Detection of shielded radionuclides from weak and poorly resolved spectra using group positive rival," *Radiation Measurements*, pp. 18–28, 2013.
[6] P. Fehlau, "Comparing a recursive digital filter with the moving-average and sequential probability-ratio detection methods for snm portal monitors," *Nuclear Science, IEEE Transactions on*, vol. 40, pp. 143–146, Apr 1993.
[7] B. Deb, J. A. F. Ross, A. Ivan, and M. Hartman, "Radioactive source estimation using a system of directional and non-directional detectors," *Nuclear Science, IEEE Transactions on*, pp. 3281–3290, Dec 2011.
[8] G. F. Knoll, *Radiation detection and measurement*. John Wiley & Sons, 2010.
[9] D. Mihalas and B. Weibel-Mihalas, *Foundations of radiation hydrodynamics*. Courier Dover Publications, 1999.
[10] C. Pahlajani, I. Poulakakis, and H. Tanner, "Decision making in sensor networks observing poisson processes," in *Control Automation (MED), 2013 21st Mediterranean Conference on*, pp. 1230–1235, June 2013.
[11] Y. Aharonov and D. Bohm, "Time in the quantum theory and the uncertainty relation for time and energy," *Physical Review*, vol. 122, no. 5, p. 1649, 1961.
[12] D. E. Archer, B. R. Beauchamp, D. A. Knapp, G. J. Mauger, M. B. Mercer, K. E. Nelson, D. C. Pletcher, V. J. Riot, and J. L. Schek, "Adaptable radiation monitoring system and method," June 20 2006. US Patent 7,064,336.
[13] F. Doty, H. Barber, F. Augustine, J. Butler, B. Apotovsky, E. Young, and W. Hamilton, "Pixellated cdznte detector arrays," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 353, no. 1, pp. 356–360, 1994.
[14] M. Barber, R. Bordoli, R. Sedgwick, and A. Tyler, "Fast atom bombardment of solids as an ion source in mass spectrometry," *Nature*, vol. 293, no. 5830, pp. 270–275, 1981.
[15] R. Pasupathy, "Generating homogeneous poisson processes," *Wiley Encyclopedia of Operations Research and Management Science*, 2010.
[16] G. F. Franklin, J. D. Powell and A. Emami-Naemi, *Feedback Control of Dynamic Systems*, Pearson, 2010.