

Algorithms for regional source localization

Sandra H. Dandach

Francesco Bullo

Abstract—In this paper we use the MAP criterion to locate a region containing a source. Sensors placed in a field of interest divide the latter into smaller regions and take measurements that are transmitted over noisy wireless channels. The sensors aggregate their data in order to identify which of these regions contains a source. We propose implementations of our algorithm that consider complete and limited communication among sensors and seek to choose the most likely hypothesis. Each hypothesis corresponds to the event that a given region contains the source. Corrupted measurements are used to calculate conditional posteriors. In the complete communication case, the sensors have full information access, whereas in the limited communication case, the sensors exchange information only with their neighbors, but each sensor tests fewer hypotheses. We prove that the algorithms asymptotically find the correct region almost surely as long as information is available from three or more sensors. We also study the geometric properties of the model that make it possible in some situations to detect the correct region with a unique sensor. Our simulations show that the performance of algorithms with complete and limited information ameliorates with decreasing noise and that in noisy environments, using fewer information from close nodes is more reliable than full information from far away nodes.

I. INTRODUCTION

A. Problem description and motivation

Source localization has assumed increasing interest, and has been the subject of study for many researchers. The general setting is that one or multiple sources lie in a bounded region C , and a group of N sensors divide C into N smaller localization regions W_i , where $i \in \{1, \dots, N\}$. They measure a received signal strength originating at a source s , the sensors try to cooperatively identify the region W_i containing s . We set the problem as a multiple hypothesis decision making problem, where hypothesis H_i is true if the source lies in the region W_i . Maximum a posteriori estimation (MAP) is used as a decision making tool. We implement the estimation technique with an all-to-all and a limited communication algorithm. The setting of the problem and the proposed solution prove to have some geometric characteristics that we derive later in the paper. If properly exploited, these characteristics imply the possibility of regional localization with a unique sensor for certain source positions. We also prove almost sure (*a.s.*) convergence of both our all-to-all and limited communication algorithms. To the best of our knowledge, none of the algorithms in the literature provide a similar convergence result.

B. Literature review

In the classical setting, a number of sensors collaborate to locate the exact position of a source. The relation between the position of a source and the received signal strength (RSS) is described in [1], [2], [3]. RSS indirectly provides

the distance between the source and a sensor. It is easy to formulate a trivial linear algorithm that permits localization from the measurements at three sensors. However, such a linear algorithm may deliver highly inaccurate estimates of the distances, even when the noise is small [4]. On the other hand, several authors [5], [6] treat localization as a nonconvex optimization problem. Gradient descent algorithms can be used to solve the maximum likelihood estimation problems. Other approaches include approximating the nonlinear nonconvex optimization problem by a linear and convex one and then proposing algorithms for the relaxed problem [7].

Following the gradient, and approximating with a linear convex problem have limitations. The gradient descent can get stuck at local minima far from the correct position, leading to the choice of wrong regions, even in the absence of noise, [8]. The authors in [7] solve the convex approximation problem using Semi-Definite Programming methods. The point of convergence of these algorithms is initialization dependent and no proof is provided on how good a solution this approximation gives. Other attempts to solve the problem is through weighted least squares approximations [6] which continues to face the problem of local minima. Authors in [5] use a method of projection onto convex sets. A necessary and sufficient condition for the convergence of this algorithm is that the source lies inside the convex hull of the sensors. The limitations of these methods motivated us to look into regional localization. We note that there are instances where the location of the region containing a source is all that is needed. A more detailed listing of the contributions of this paper is presented below.

C. Contributions

This paper presents the source localization problem in a setting and formulation that to the best of our knowledge are new. We present algorithms based on all-to-all and limited communication that require only the computation of integrals and therefore present a less computationally exhaustive alternative to the current solutions to the localization problem. We also show that as the noise decreases, regional localization can be accomplished with a unique sensor for certain source positions. We show through an asymptotic analysis that choosing Voronoi partitions as localization regions achieves zero probability of error in the two sensors case. The most important advantage of our formulation is that we are able to demonstrate the convergence of our algorithms. We provide the proof of *a.s.* convergence of our algorithms, a step that tends to be missing in all of the work presented earlier. Finally, the limited communication algorithm is promising for the localization problems involving multiple sources.

D. Paper organization

The paper proceeds by a problem formulation and an explanation of our proposed solution in Section II. In Section III we derive some asymptotic geometric properties of

The authors are with the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, Santa Barbara, CA 93106, USA, {sandra,bullo}@engineering.ucsb.edu. This work was supported in part by Award AFOSR FA9550-07-1-0528.

the MAP algorithm when applied to our setting. Section IV introduces the implementation of the algorithms. The analytical proof of almost sure (*a.s.*) convergence of our algorithms is presented in Section V. Section VI shows our simulation results and we conclude in Section VII.

II. PROBLEM FORMULATION

Consider a compact connected environment $C \subset \mathbb{R}^2$. Suppose there are N disjoint regions W_i , such that $\cup_{i=1}^N W_i = C$. Suppose also that there are N sensors placed at $x_i \in W_i$ and that the source located at an unknown location $s \in C$, transmits a signal whose power undergoes lognormal shadowing described below.

The average power loss for an arbitrary Transmitter-Receiver separation is expressed as a function of distance by using a path loss exponent $\beta > 2$. The power loss is proportional to a power of the distance between the transmitter and the receiver. For a thorough description of signal attenuation models over communication channels, we refer the reader to [1], [2], [7]. For reasons to be explained shortly, we work with a slight modification of the traditional model used in [1], [2], [7]. This model for the received power at a sensor i is, $P_{r_i} = \frac{P d_0}{d_0 + \|x_i - s\|^\beta}$, where β indicates the rate at which the power loss increases with distance. d_0 is a nominal distance chosen such that the received power in the vicinity of the source is almost equal to P , the transmitted power at the source. Note that while this model gets rid of the singularity at the source, it converges to the same behavior as the classical model used in communication literature $P_{r_i} = \frac{P_1}{\|x_i - s\|^\beta}$, when the distance $\|x_i - s\|$ is large. Here P_1 is the power received at a unit distance from the source. Taking noise into account in our model, the received power satisfies

$$\ln P_{r_i} = \ln(P d_0) - \ln(d_0 + \|x_i - s\|^\beta) + n_i, \quad (1)$$

where n_i are zero mean, independent and identically distributed (i.i.d) white gaussian noise with variance σ^2 , each associated with a sensor i . The joint probability density function of the $\ln P_r = [\ln P_{r_1}, \dots, \ln P_{r_N}]^T$, conditioned on the hypothesis that the source is at a point $y \in C$ is given by

$$p(\ln P_{r_1}, \dots, \ln P_{r_N} | y) = \frac{1}{(2\pi\sigma^2)^{N/2}} \cdot \exp\left(-\frac{\sum_{i=1}^N \left(\ln P_{r_i} - \ln\left(\frac{P d_0}{d_0 + \|x_i - y\|^\beta}\right)\right)^2}{2\sigma^2}\right). \quad (2)$$

Solving for the exact position of the source requires solving for \hat{y} that will maximize the likelihood of having the received observation which becomes the problem of solving for,

$$\hat{y} = \arg \min_y \sum_{i=1}^N \left(\ln P_{r_i} - \ln\left(\frac{P d_0}{d_0 + \|x_i - y\|^\beta}\right)\right)^2.$$

This is a nonlinear nonconvex optimization problem. Attempts to solve it or approximate its solution are a topic of great interest. In this paper we look for a regional localization, so the conditioning on the exact position y in (2) is replaced by a regional conditioning. Before we elaborate on this point, we introduce the information exchanged among

the sensors for the all-to-all and the limited communication algorithms.

As sensors share information, they collaborate to decide which of the N regions contains the source. The information exchanged between any two communicating sensors are: the position of the sensors, the localization regions associated with each sensor and the logarithms of the received powers (corrupted with log-normal noise). Sensors can share information as soon as they make a measurement. Alternatively when the noise level in the communication channel is known to be high, it is possible for each sensor to average a set of repeated measurements and transmit the averaged logarithm of the received power. Averaging helps decrease the noise variance, and therefore as we expect and will show later, improves the performance. We start by introducing the case of one noisy measurement per sensor.

A. Posterior density with a single noisy measurement

Since we do not know where the source is, we make a worst case assumption on the knowledge of its position s . Specifically, we assume that the density of s obeys:

$$p(s) = \begin{cases} 1/A, & \text{if } s \in C, \\ 0, & \text{otherwise.} \end{cases}$$

Here A is the sum of all the areas A_j of W_j , with $j \in \{1, \dots, N\}$. We need to derive the probability density conditioned on each hypothesis.

Lemma II.1 (Regional conditional density) Let

$z = [\ln P_{r_1}, \dots, \ln P_{r_N}]^T$, and note that $P(y \in W_j) = P(H_j) = \frac{A_j}{A}$, then

$$p(z|y \in W_j) = \frac{1}{A_j} \int_{W_j} p(z|y) dy. \quad (3)$$

Proof: We compute

$$\begin{aligned} p(z|y \in W_j) &= \frac{d \text{Prob}(Z \leq z, y \in W_j)}{P(y \in W_j)} \\ &= A \frac{d \int_{-\infty}^z \int_{W_j} p(z|y) p(y) dy dz}{A_j} \\ &= A \frac{d \int_{-\infty}^z \int_{W_j} (p(z|y)/A) dy dz}{A_j} = \frac{\int_{W_j} p(z|y) dy}{A_j}. \end{aligned}$$

From (2) and (3), we obtain

$$\begin{aligned} p(\ln P_{r_1}, \dots, \ln P_{r_N} | H_j) P(H_j) &= \frac{1}{A} \int_{W_j} \frac{1}{(2\pi\sigma^2)^{N/2}} \\ &\cdot \exp\left(-\frac{\sum_{i=1}^N \left(\ln P_{r_i} - \ln\left(\frac{P d_0}{d_0 + \|x_i - y\|^\beta}\right)\right)^2}{2\sigma^2}\right) dy. \quad (4) \end{aligned}$$

Similarly, when measurements from only one sensor are studied, the conditional probability density is derived to be

$$\begin{aligned} p(\ln P_{r_i} | H_j) P(H_j) &= \frac{1}{A} \int_{W_j} \frac{1}{(2\pi\sigma^2)^{1/2}} \\ &\cdot \exp\left(-\frac{\left(\ln P_{r_i} - \ln\left(\frac{P d_0}{d_0 + \|x_i - y\|^\beta}\right)\right)^2}{2\sigma^2}\right) dy. \quad (5) \end{aligned}$$

B. Posterior density with aggregated noisy measurements

In this setting each sensor is allowed to take k repeated noisy measurements. Noise independence is assumed between sensors and between different samples times for each sensor. Then defining

$$\mathbf{P}_{r_i}(k) = \sum_{l=1}^k \frac{P_{r_i}(l)}{k}, \quad (6)$$

the variance of the noise becomes $\sigma^2(k) = \frac{\sigma^2}{k}$. The conditional probability in (5) then becomes

$$p(\ln \mathbf{P}_{r_i}(k) | H_j) P(H_j) = \frac{1}{A} \int_{W_j} \frac{1}{(2\pi\sigma^2(k))^{1/2}} \cdot \exp\left(-\frac{\left(\ln \mathbf{P}_{r_i}(k) - \ln\left(\frac{Pd_0}{d_0 + \|x_i - y\|^\beta}\right)\right)^2}{2\sigma^2(k)}\right) dy,$$

and the joint conditional probability becomes:

$$p(\ln \mathbf{P}_{r_1}(k), \dots, \ln \mathbf{P}_{r_N}(k) | H_i) P(H_i) = \frac{1}{A} \int_{W_i} dy \prod_{j=1}^N \frac{1}{(2\pi\sigma^2(k))^{1/2}} \exp\left(-\frac{\left(\ln \mathbf{P}_{r_j}(k) - \ln\left(\frac{Pd_0}{d_0 + \|x_j - y\|^\beta}\right)\right)^2}{2\sigma^2(k)}\right).$$

Note that, as $k \rightarrow \infty$, the noise variance approaches zero, and the probability density approaches a delta function.

Remark II.2 Let δ be the Dirac delta function. In the infinite measurement case, $\lim_{k \rightarrow \infty} \sigma^2(k) = 0$, and the probability density satisfies

$$\begin{aligned} p(\ln \mathbf{P}_{r_i} | y) &= \lim_{k \rightarrow \infty} \frac{1}{(2\pi\sigma(k)^2)^{1/2}} \exp\left(-\frac{\left(\ln \mathbf{P}_{r_i} - \ln\left(\frac{Pd_0}{d_0 + \|x_i - y\|^\beta}\right)\right)^2}{2\sigma^2(k)}\right) \\ &= \delta\left(\ln \mathbf{P}_{r_i} - \ln \frac{Pd_0}{d_0 + \|x_i - y\|^\beta}\right). \quad \square \end{aligned}$$

Remark II.3 In the sequel, for notational simplicity we will treat the aggregated measurement case as if it were identical to the single measurement case with the caveat that the variance goes to zero. \square

C. All-to-all information MAP estimation

In the all-to-all communication (A2A) case, full information is available. Using the conditional probability in (4) MAP selects the hypothesis H_{i^*} according to

$$i^* = \arg \max_i p(\ln P_{r_1}, \dots, \ln P_{r_n} | H_i) P(H_i). \quad (7)$$

Per Remark II.3, this selection scheme applies to both the single and the aggregated measurement cases.

Before we proceed to deriving the results in the next section, we introduce the definition of the Voronoi diagrams.

Definition II.4 (Voronoi Diagrams) Given N sensors located at positions $\{x_1, \dots, x_N\} \in C$, we define the Voronoi diagram associated with the i th sensor, as follows

$$V_i = \{x \in C : \|x - x_i\| \leq \|x - x_j\|, \forall j \neq i\}.$$

III. PRELIMINARY PROPERTIES OF REGIONAL LOCALIZATION FOR ONE AND TWO SENSORS

In this section we derive certain geometric properties of MAP estimation as $k \rightarrow \infty$ in (6). These geometric properties allow us to conclude the following two results. First, for certain source locations, a single sensor suffices to asymptotically detect the correct hypothesis. Second, for the asymptotic detection problem with two sensors, the selection of Voronoi partitions as localization regions leads to exact localization. These 2 results should be viewed against the fact that, even in the noise-free case, at least 3 non-collinear sensors are needed for exact localization. In this section we conduct a large sample analysis to prove an interesting geometric interpretation of the conditional probability densities. This analysis recognizes that when $k \rightarrow \infty$ in (6), the Gaussian density approaches a Dirac delta function. Before we state the lemma that captures this property, we mention a basic property of the Dirac delta function [9].

Lemma III.1 (On the Dirac delta function) If $g : \mathbb{R} \rightarrow \mathbb{R}$ is differentiable and vanishes at positions x_γ , $\gamma \in \Gamma$, then

$$\delta[g(x)] = \sum_{\gamma \in \Gamma} \frac{\delta(x - x_\gamma)}{|g'(x_\gamma)|}. \quad (8)$$

In keeping with Remark II.3, consider the situation where

$$\begin{aligned} p(\ln P_{r_i} | y \in W_j) P(y \in W_j) &= \frac{1}{A} \int_{W_j} \delta\left(\ln P_{r_i} - \ln \frac{Pd_0}{d_0 + \|x_i - y\|^\beta}\right) dy. \end{aligned}$$

Define the circle $C(r, x_i) = \{y \in \mathbb{R}^2 \mid \|y - x_i\| = r\}$ and denote its intersection with the region W_j by $S(W_j, r, x_i) = C(r, x_i) \cap W_j$. Clearly, this intersection set $S(W_j, r, x_i)$ is the union of certain arcs of $C(r, x_i)$. Define $\theta(W_j, r, x_i)$ to be the sum of the angles subtended by these arcs. Call $H_j = (y \in W_j)$. If we let $y = [y_1, y_2]^T$ and define

$$\begin{aligned} f(y_1, y_2, P_{r_i}) &= \ln P_{r_i} - \ln Pd_0 \\ &\quad + \ln(d_0 + ((x_{i1} - y_1)^2 + (x_{i2} - y_2)^2)^{\beta/2}), \end{aligned}$$

then

$$p(\ln P_{r_i} | H_j) P(H_j) = \frac{1}{A} \int_{W_j} \delta(f(y_1, y_2, P_{r_i})) dy_2 dy_1.$$

We are now ready for the following lemma.

Lemma III.2 (The arc-length property) Given a region W_j , the conditional probability density satisfies

$$\begin{aligned} p(\ln P_{r_i} | y \in W_j) P(y \in W_j) &= \frac{1}{A} \int_{W_j} \delta\left(\ln P_{r_i} - \frac{Pd_0}{d_0 + \|x_i - y\|^\beta}\right) dy_2 dy_1, \end{aligned}$$

and, if we let $r_i = \left(\frac{P}{P_{r_i}} - 1\right)^\beta$, then

$$p(\ln P_{r_i} | y \in W_j) P(y \in W_j) = \frac{d_0 + r_i^\beta}{A\beta r_i^{\beta-2}} \theta(W_j, r_i, x_i).$$

The proof of the lemma is given in the Appendix. This lemma can be interpreted as follows. Asymptotically, P_{r_i}

directly provides the circle of radius r_i centered on x_i on which the source lies. $\theta(W_j, r_i, x_i)$ is simply the angle subtended by the intersection of this circle with W_j . The quantity $p(\ln P_{r_i} | y \in W_j) P(y \in W_j)$ that is used in MAP is proportional to this angle. We describe further the significance of this result after Lemma III.3. In the following lemma, we show that having Voronoi partitions as well as the MAP estimation algorithm, make the probability of error zero in the two sensors case.

Lemma III.3 (Optimality of Voronoi for 2 sensors)

Consider two points x_1 and x_2 in $C \subset \mathbb{R}^2$. Let V_1 and V_2 be the Voronoi diagrams associated with x_1 and x_2 . Take $s \in C$ and let $r_1 = \|s - x_1\|$ and $r_2 = \|s - x_2\|$. Then as k in (6) tends to infinity, MAP localization algorithm finds the region containing s , with zero probability of error with only two sensors.

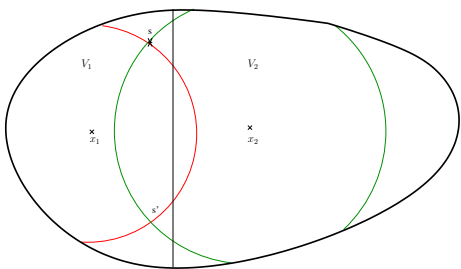


Fig. 1. This figure shows two nodes x_1 and x_2 , with a source $s \in V_1$.

Proof: The proof of this lemma follows the same principle as the proof of Lemma III.2. We skip the details for lack of space, but we point out the differences below. In figure 1, the source s is indicated as one of the two points of intersection of the red and green circle. The term h defined in (14) becomes $h(y_1, P_{r_1}, x_1, P_{r_2}, x_2) = \int_{H(a, W_j)} \delta(f(a, y_2, P_{r_1}, x_1)) \cdot \delta(f(a, y_2, P_{r_2}, x_2)) dy_2$, here P_{r_1} and P_{r_2} are the received powers by the sensors. The integrals are non zero only at the intersections of sets $C_1 \cup C'_1$ and $C_2 \cup C'_2$ as defined in the appendix. The intersection of these two sets, is the intersection of the circles in Figure 1. The definition of Voronoi (Def. II.4) implies in the two sensors case that if $s' \in C$, then $s \in V_1 \implies s' \in V_1$. In fact $s \in V_1 \implies \|x_1 - s\| \leq \|x_2 - s\| \implies \|x_1 - s'\| \leq \|x_2 - s'\| \implies s' \in V_1$. The other case to consider is $s' \notin C$. In this case s is the unique point of intersection of the circles in C . Both ways h is non-zero only in the correct region. ■

The two lemmas presented in this section have interesting implications. Lemma III.2 implies that, for certain source locations and as the noise becomes smaller, the MAP estimation algorithm can determine the correct region containing the source with only one sensor. That is true when the circle centered at a sensor location with radius r_i is included in the region W_j . Lemma III.3 on the other hand, gives one example where the selection of Voronoi partitions as localization regions makes it possible to locate the source with only two sensors. This would not have been the case with two sensors with general convex regions.

In general, the noise will not be vanishing and the decision needs to be made with a finite number of measurements.

We will prove in Section V that the algorithms presented in Section IV below converge almost surely as $k \rightarrow \infty$ in (6).

IV. DECISION MAKING ALGORITHMS WITH ALL-TO-ALL OR LIMITED INFORMATION

In this section we present two algorithms based on the MAP estimation scheme. The regions can take any shape as long as they are compact. In the limited communication case, sensors can only talk to their neighbors. We assume there is a communication graph that describes the information exchange among robots. We consider two cases: the all-to-all (A2A) communication case and the limited communication case. In the A2A we adopt the complete undirected communication graph. In the limited communication case we make the following degree assumption: each node has at least two neighbors, i.e., each node appears in at least two edges.

A. All-to-all communication

In this subsection we present the all-to-all communication algorithm (Algorithm #1), where we apply the classical MAP estimation described in Section II-C on the complete network.

Algorithm #1: All-to-all communication MAP

Network: nodes $\{1, \dots, N\}$ with complete communication graph
State of sensor i is $w_i := \{\ln P_{r_i}, x_i, W_i\}$

Sensor i executes

- 1: **FOR** all $j \in \{1, \dots, N\} \setminus \{i\}$
- 2: transmit state w_i and receive state w_j
- 3: calculate $\theta_j := p(\ln P_{r_1}, \dots, \ln P_{r_N} | H_j) P(H_j)$
- 4: find $j^* := \arg \max_{j \in \{1, \dots, N\}} \theta_j$
- 5: **RETURN:** decision j^*

Assumption IV.1 (A2A connectivity and noncollinearity)

We assume that the graph is complete. That is all nodes can communicate with each other. We also assume that at least three sensors in the graph are non-collinear.

B. Limited communication

The algorithm in this subsection has two major differences from the all-to-all communication algorithm. First, the sensors deal with information coming from a different set of sensors and a different number of sensors. Second each sensor decides about its neighborhood and the complement of the latter.

In the limited communication algorithm, each sensor acquires data from its neighbors and calculates a joint conditional density (4). Each sensor then applies MAP to choose the most likely hypothesis. In Algorithm #2 each sensor adds the hypothesis of the source being outside its neighborhood, computes the corresponding conditional density and compares it to the densities corresponding to neighboring regions. Let $\mathcal{N}_i = \{\{j \in \{1, \dots, N\} \mid j \text{ is a neighbor of } i\} \cup \{i\}\}$ and let $N_i = |\mathcal{N}_i|$ be its cardinality. We described the ‘‘source outside neighborhood’’ hypothesis as hypothesis number 0 (H_0). We also define $\ln P_{r_{\mathcal{N}_i}}$ to be the vector composed of all measurements $\ln P_{r_i}$ where $i \in \mathcal{N}_i$. Finally, we mention that given an event $H_i = x \in W_i$, the complement event is defined by $\bar{H}_i = x \in W_i^C$. Computing the density in the complement of the neighborhood requires only the addition

of information about the total region. In fact by the total probability theorem we note that

$$\begin{aligned}
p(\ln P_{r_j}) &= p(\ln P_{r_j} | y \in C) P(y \in C) \\
&= \sum_{i \in \mathcal{N}_j} p(\ln P_{r_j} | H_i) P(H_i) + p(\ln P_{r_j} | \bigcup_{i \in \mathcal{N}_j} \overline{H_i}) P(\bigcup_{i \in \mathcal{N}_j} \overline{H_i}) \\
&\implies p(\ln P_{r_j} | H_0) P(H_0) = p(\ln P_{r_j} | \bigcup_{i \in \mathcal{N}_j} \overline{H_i}) P(\bigcup_{i \in \mathcal{N}_j} \overline{H_i}) \\
&= p(\ln P_{r_j} | y \in C) P(y \in C) - \sum_{i \in \mathcal{N}_j} p(\ln P_{r_j} | H_i) P(H_i)
\end{aligned}$$

Algorithm #2: Limited communication MAP

Network: nodes $\{1, \dots, N\}$ with arbitrary communication graph
State of sensor i is $w_i := \{\ln P_{r_i}, x_i, W_i\}$

Sensor i executes

- 1: For all $j \in \mathcal{N}_i$
 - 2: transmit state w_i and receive state w_j
 - 3: calculate $\theta_{i,j} := p(\ln P_{r_{\mathcal{N}_i}} | H_j) P(H_j)$
 - 4: calculate $\theta_{i,0} = p(\ln P_{r_{\mathcal{N}_i}} | H_0) P(H_0)$
 - 5: find $j^* := \arg \max_{i \in \mathcal{N}_i \cup \{0\}} \theta_{i,j}$
 - 6: **Return:** decision j^*
-

Assumption IV.2 (Limited conn. and non-collinearity)

We assume that all nodes can communicate to their neighbors. We also assume that at least three sensors are non-collinear in each neighborhood.

V. CONVERGENCE OF ALGORITHMS

In this section we prove that the algorithms presented above give the correct decision *w.p.1* with three or more non-collinear sensors. In keeping with Remark II.3, we examine (4) as $\sigma \rightarrow 0^+$. We start by stating the following result.

Lemma V.1 (Property of non-collinear sensors) For $d_0 > 0$ and $\beta > 0$, given a source $s \in \mathbb{R}^2$ and three non-collinear sensors x_1, x_2 and $x_3 \in \mathbb{R}^2$, define the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by $f(z) = \sum_{i=1}^3 \left(\ln \frac{d_0 + \|z - x_i\|^\beta}{d_0 + \|s - x_i\|^\beta} \right)^2$. The function $f(z)$ vanishes only if and only if $z = s$.

Proof: In fact, it is easy to check that the sum is zero at $z = s$. Uniqueness of this solution is verified by noting that the sum of the square terms is zero only if all the summands are zero. For that to be true one needs to find $z = (x, y)$ such that $((x - x_{i1})^2 + (y - x_{i2})^2)^{\beta/2} = ((s_1 - x_{i1})^2 + (s_2 - x_{i2})^2)^{\beta/2} \doteq (r_i^2)^{\beta/2}$, for $i \in \{1, 2, 3\}$. Expanding and subtracting, we obtain

$$\begin{aligned}
&\begin{bmatrix} -2(x_1 - x_2) & -2(y_1 - y_2) \\ -2(x_1 - x_3) & -2(y_1 - y_3) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \begin{bmatrix} (r_1^2 - d_1^2) - (r_2^2 - d_2^2) \\ (r_3^2 - d_3^2) - (r_1^2 - d_1^2) \end{bmatrix}, \quad (9)
\end{aligned}$$

where $d_0^2 = x^2 + y^2$, and $d_i^2 = x_i^2 + y_i^2$. Equation (9) has a unique solution if and only if the determinant of the matrix is non zero, i.e., the three points are non-collinear. ■

In fact given non-collinear $x_i, i \in \{1, \dots, N\}$ with $N \geq 3$, a compact region W_j and a source $s \notin W_j$, there always exists

$$D_j = \min_{y \in W_j} \|s - y\| = \text{dist}(W_j, s),$$

such that $D_j > 0$. Also for a compact regions W_j , there exists $U_j > 0$ such that

$$U_j \geq \max_{y \in W_j, i \in \{1, \dots, N\}} \left| \ln \frac{d_0 + \|y - x_i\|^\beta}{d_0 + \|s - x_i\|^\beta} \right|. \quad (10)$$

Because of the non-collinearity and the fact that $D_j > 0$, there exists an $L_j > 0$ such that

$$L_j \leq \min_{y \in W_j} \sum_{i=1}^N \left(\ln \frac{d_0 + \|y - x_i\|^\beta}{d_0 + \|s - x_i\|^\beta} \right)^2. \quad (11)$$

This follows from the lemma above (i.e., from the fact that the sum has a unique global minimum at s). Define

$$\eta_j = \sqrt{U_j^2 + \frac{L_j}{\alpha N}} - U_j, \quad (12)$$

where $\alpha > 1$, then we have the following result.

Lemma V.2 Consider L_j, U_j and η_j as defined in (10), (11) and (12). Suppose the source s is not in region W_j , and $|n_i| \leq \eta_j$ for all $i \in \{1, \dots, N\}$. Then

$$\frac{\alpha - 1}{\alpha} L_j \leq \min_{y \in W_j} \sum_{i=1}^N \left(\ln \frac{d_0 + \|y - x_i\|^\beta}{d_0 + \|s - x_i\|^\beta} + n_i \right)^2. \quad (13)$$

Proof: In fact, the sum in (13) satisfies:

$$\begin{aligned}
&\sum_{i=1}^N \left(\ln \frac{d_0 + \|y - x_i\|^\beta}{d_0 + \|s - x_i\|^\beta} + n_i \right)^2 \\
&= \sum_{i=1}^N \left(\ln \frac{d_0 + \|y - x_i\|^\beta}{d_0 + \|s - x_i\|^\beta} \right)^2 \\
&\quad + 2 \sum_{i=1}^N \left(\ln \frac{d_0 + \|y - x_i\|^\beta}{d_0 + \|s - x_i\|^\beta} \right) n_i + \sum_{i=1}^N n_i^2 \\
&\geq L_j - 2U_j N \eta_j - N \eta_j^2 = L_j + 2U_j^2 N \\
&\quad - 2U_j N \sqrt{U_j^2 + \frac{L_j}{\alpha N}} - N \left(U_j^2 + \frac{L_j}{\alpha N} \right) - N U_j^2 \\
&\quad + 2N U_j \sqrt{U_j^2 + \frac{L_j}{\alpha N}} = \frac{\alpha - 1}{\alpha} L_j.
\end{aligned}$$

For simplicity of notation, we choose $\alpha = 2$ from here on. Next, we introduce the final intermediate lemma before the main results of this section.

Lemma V.3 (Upper bound for wrong hypothesis) Given L_j, U_j and η_j as defined in (10), (11) and (12), define

$$\begin{aligned}
I_j &= p(\ln P_{r_1}, \dots, \ln P_{r_N} | H_j) P(H_j) = \frac{1}{A(2\pi\sigma^2)^{N/2}} \\
&\cdot \int_{W_j} \exp\left(- \frac{\sum_{i=1}^N \left(\ln P_{r_i} - \ln \frac{P d_0}{d_0 + \|y - x_i\|^\beta} \right)^2}{2\sigma^2} \right) dy.
\end{aligned}$$

If $|n_i| \leq \eta_j$ for $i \in \{1, \dots, N\}$, then for $j \in \{1, \dots, N\}$

$$I_j \leq \frac{A_j \exp(-L_j/4\sigma^2)}{A(2\pi\sigma^2)^{N/2}}.$$

Proof: Because of the equality

$$\ln P_{r_i} - \ln \frac{Pd_0}{d_0 + \|y - x_i\|^\beta} = \ln \frac{d_0 + \|y - x_i\|^\beta}{d_0 + \|s - x_i\|^\beta} + n_i,$$

the result directly follows from Lemma V.2 and from the fact that the surface integral of a function f is upper bounded by the surface integral of a constant function g , where g takes the maximum value of f . ■

We are now ready for the convergence theorem. As usual, we define the Q -function $Q : \mathbb{R} \rightarrow \mathbb{R}_{>0}$ by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \exp(-y^2/2) dy.$$

Theorem V.4 (Elimination property of wrong hypothesis)

Consider x_i non-collinear sensors, $i \in \{1, \dots, N\}$ with $N \geq 3$. Let σ be the noise variance. Given a source $s \notin W_j$, then we have

$$\text{Prob} \left[p(\ln P_{r_1}, \dots, \ln P_{r_N} | H_j) P(H_j) \leq \epsilon_j(\sigma) \right] \geq \mu_j(\sigma),$$

where

$$\epsilon_j(\sigma) = \frac{A_j \exp(-L_j/4\sigma^2)}{A(2\pi\sigma^2)^{N/2}}, \quad \mu_j(\sigma) = (1 - 2Q(\eta_j/\sigma))^N.$$

Furthermore, as $\sigma \rightarrow 0^+$, we have $\epsilon_j(\sigma) \rightarrow 0^+$ and $\mu_j(\sigma) \rightarrow 1^-$.

Proof: From Lemmas V.2 and V.3, we have that

$$\begin{aligned} & \text{Prob} [p(\ln P_{r_1}, \dots, \ln P_{r_N} | H_j) P(H_j) \leq \epsilon_j(\sigma)] \\ & \geq \text{Prob} [[n_1, \dots, n_N]^T \in [-\eta_j, \eta_j]^N] \\ & = \prod_{i=1}^N \left(\frac{1}{2} - \text{Prob}[n_i > \eta_j] + \frac{1}{2} - \text{Prob}[n_i < -\eta_j] \right) \\ & = (1 - 2Q(\eta_j/\sigma))^N. \end{aligned}$$

The first inequality comes from the fact that Lemmas V.2 and V.3 hold whenever all $|n_i| \leq \eta_j$. The proofs of the two limits of ϵ_j and μ_j are immediate. ■

This theorem states that, as $\sigma \rightarrow 0^+$, the probability that the joint density function $p(\ln P_{r_1}, \dots, \ln P_{r_N} | H_j) P(H_j)$ takes an arbitrarily small value goes arbitrarily close to 1 when H_j is not the correct hypothesis. This is so as $Q(x) \rightarrow 0$ as $x \rightarrow \infty$. To complement the Theorem V.4, we prove below that for the *correct* hypothesis, the probability density will be lower bounded by a positive term *w.p.1*.

Theorem V.5 (Strict positivity for correct hypothesis)

Consider x_i non-collinear sensors, $i \in \{1, \dots, N\}$ with $N \geq 3$. Let σ be the noise variance. If $s \in W_i$, then we have

$$\text{Prob} [p(\ln P_{r_1}, \dots, \ln P_{r_N} | H_i) P(H_i) \geq \Psi(\sigma)] \geq \Omega(\sigma),$$

where

$$\Psi(\sigma) = p(\ln P_{r_1}, \dots, \ln P_{r_N}) - \sum_{\substack{j=1, \dots, N \\ j \neq i}} \frac{A_j \exp(-L_j/4\sigma^2)}{A(2\pi\sigma^2)^{N/2}},$$

$$\Omega(\sigma) = \prod_{\substack{j=1, \dots, N \\ j \neq i}} \mu_j(\sigma) = \prod_{\substack{j=1, \dots, N \\ j \neq i}} (1 - 2Q(\eta_j/\sigma))^N.$$

Furthermore, as $\sigma \rightarrow 0^+$, we have $\Psi(\sigma) \rightarrow p(\ln P_{r_1}, \dots, \ln P_{r_N}) > 0$ and $\Omega(\sigma) \rightarrow 1^-$.

Proof: The proof of this theorem follows directly from Theorem V.4 and the total probability theorem. Call $z = [\ln P_{r_1}, \dots, \ln P_{r_N}]^T$. We know from the total probability theorem that

$$\begin{aligned} p(z) &= \sum_{j=1}^N p(z|H_j) P(H_j) \\ &= p(z|H_i) P(H_i) + \sum_{\substack{j=1, \dots, N \\ j \neq i}} p(z|H_j) P(H_j) \end{aligned}$$

and, in turn, that

$$p(z|H_i) P(H_i) = p(z) - \sum_{\substack{j=1, \dots, N \\ j \neq i}} p(z|H_j) P(H_j).$$

From Theorem V.4

$$\begin{aligned} & \text{Prob} \left[p(z|H_i) P(H_i) \geq p(z) - \sum_{\substack{j=1, \dots, N \\ j \neq i}} \epsilon_j(\sigma) \right] \\ & \geq \prod_{\substack{j=1, \dots, N \\ j \neq i}} \text{Prob} \left[p(z|H_j) P(H_j) \leq \epsilon_j(\sigma) \right] \\ & \geq \prod_{\substack{j=1, \dots, N \\ j \neq i}} \mu_j(\sigma) \end{aligned}$$

As $\sigma \rightarrow 0^+$, $\Psi(\sigma) \rightarrow p(z)$ and $\Omega(\sigma) \rightarrow 1^-$. ■

Theorem V.5 complements Theorem V.4 in that it shows that as $\sigma \rightarrow 0^+$, the probability density conditioned on the correct hypothesis is lower bounded by a strictly positive term. This event happens asymptotically with probability 1.

Under Assumption IV.1, as MAP follows (7), Theorems V.5 and V.4 complete the proof of *a.s* convergence of the all-to-all communication MAP algorithm. Similarly for the limited communication, under Assumption IV.2, MAP estimation converges almost surely when applied to regional localization.

VI. SIMULATIONS

In this section we show simulation results illustrating the type of decision obtained by our algorithms. We also show a comparison plot between the all-to-all and the limited information algorithms. The plot in Figure 2 shows a correct

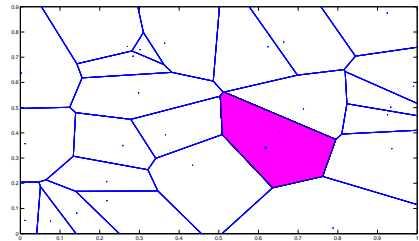


Fig. 2. This figure shows the all-to-all communication algorithm with 25 sensors and one source whose region was correctly detected.

detection of the source. The shaded region corresponds to the one detected by the algorithm, the source is shown as a star and the sensors as the dots. Figure 3 shows the results

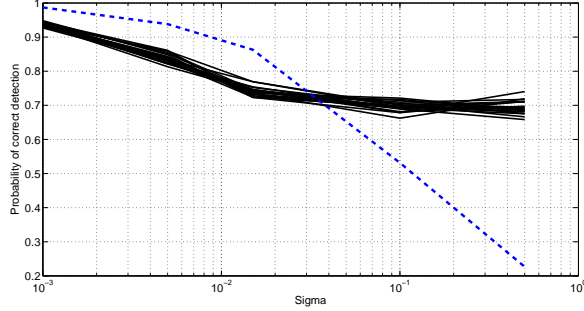


Fig. 3. This figure shows the probability of correct decision making of the all-to-all and limited communication cases over 1000 runs, with 20 sensors. The stand alone dotted blue curve corresponds to the A2A communication algorithm, and the curves bunched up together in solid black correspond to the limited communication algorithm.

obtained from batches of 1000 runs, with $N = 20$ sensors. The plots are obtained as follows: If d_j is the index of the region chosen by sensor j , we compare it to the correct index d^* . An error is counted whenever the two indices are not equal. In the all-to-all communication algorithm, $d_j = d$ is the same for all $j \in \{1, \dots, N\}$. In the limited communication case, a decision d_j is made about the sensors in the neighborhood of sensor j , and an additional region containing the remaining part of the field C is added to the localization regions being tested. d_j is compared to d^* to check if the correct decision was made. An error is counted every time sensor j gives a wrong decision over the 1000 runs. Errors are counted for each sensor separately. The curves that are in solid black and bunched up together in Figure 3, represent the limited communication decisions. The stand alone dotted blue curve represents the all-to-all communication decision. Figure 3 shows the probability of making a correct decision as σ increases. It is not surprising that as the noise variance increases the probability of correct decision making decreases, but it is interesting to see that as the noise becomes larger, an algorithm where sensors communicate to their neighbors only performs better than an all to all communication algorithm. One interpretation is that in the limited communication case a decision is made about closer regions than in the other case.

VII. CONCLUSION

In this paper we have presented an all-to-all and a limited communication algorithm based on MAP that succeed in identifying a region that contains a source. We have also presented an asymptotic analysis and derived some geometric properties of our algorithms. Those properties had the implication that for certain source positions, it is possible to solve the regional localization problem with a unique sensor. We showed that choosing Voronoi partitions for localization regions asymptotically achieves zero probability of error in the two sensors case. We were also able to prove that when readings from three or more non-collinear sensors are available, the algorithms choose the correct region almost surely. As an extension to this work we are studying how to optimally position the sensors, and choose the localization

regions, so that we minimize the probability of error of our algorithms.

REFERENCES

- [1] T. S. Rappoport, *Wireless Communications: Principles and Practice*. New Jersey: Prentice Hall, 1996.
- [2] J. Proakis and M. Salehi, *Communication Systems Engineering*. Prentice Hall, 2001.
- [3] A. H. Sayed, A. Tarighat, and Khajehnouri, "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Processing Magazine*, pp. 24–40, 2005.
- [4] M. Cao, B. Anderson, and A. Morse, "Sensor network localization with imprecise distance measurements," *Systems & Control Letters*, pp. 87–93, 2006.
- [5] A. Hero and D. Blatt, "Sensor network source localization via projection onto convex sets (pcos)," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 689–692, 2005.
- [6] M. G. Rabbat and R. D. Nowak, "Decentralized source localization and tracking," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 921–924, 2004.
- [7] C. Meng, Z. Ding, and S. Dasgupta, "A semidefinite programming approach to source localization in wireless network," *IEEE Signal Processing Letters*, pp. 389–404, 2008.
- [8] G. Mao, B. Fidan, and B. D. O. Anderson, "Wireless sensor network localization techniques," *Computer Network*, vol. 51, no. 10, pp. 2529–2553, 2007.
- [9] K. T. Hecht, *Quantum Mechanics*. USA: Springer, 1996.
- [10] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. San Diego: Academic Press, 1994.

APPENDIX

Proof: [Proof of Lemma III.2] Let

$$H(a, W_j) = (y_2 \in \mathbb{R} \text{ such that given } y_1 = a, [a, y_2] \in W_j)$$

$$\begin{aligned} h(y_1, P_{r_i}, x_i) &= \int_{H(a, W_j)} \delta \left(\ln P_{r_i} - \ln \frac{P d_0}{d_0 + \|x_i - y\|^\beta} \right) dy_2 \\ &= \int_{H(a, W_j)} \delta(f(a, y_2, P_{r_i}, x_i)) dy_2. \end{aligned} \quad (14)$$

Since

$$\begin{aligned} \frac{d}{dy_2} f(y_1, y_2, P_{r_i}, x_i) &= f'(y_1, y_2, P_{r_i}, x_i) \\ &= \frac{\beta}{2} \cdot 2 \cdot (-1) \cdot (x_{i_2} - y_2) \frac{((x_{i_1} - y_1)^2 + (x_{i_2} - y_2)^2)^{\frac{\beta}{2}-1}}{d_0 + ((x_{i_1} - y_1)^2 + (x_{i_2} - y_2)^2)^{\frac{\beta}{2}}} \end{aligned}$$

If we fix $y_1 = a$, we can solve for $y_2(a)$ such that, $f(a, y_2(a), P_{r_i}, x_i) = 0$. In fact

$$\begin{aligned} f(a, y_2(a), P_{r_i}, x_i) &= 0 \\ \Leftrightarrow \ln P_{r_i} - \ln \frac{P d_0}{d_0 + ((x_{i_1} - a)^2 + (x_{i_2} - y_2(a))^2)^{\frac{\beta}{2}}} &= 0 \\ \Leftrightarrow (x_{i_1} - a)^2 + (x_{i_2} - y_2(a))^2 &= \left(\frac{P - P_{r_i} d_0}{P_{r_i}} \right)^{\frac{2}{\beta}} = r_i^2, \end{aligned} \quad (15)$$

where $r_i = \left(\left(\frac{P}{P_{r_i}} - 1 \right) d_0 \right)^{\frac{1}{\beta}}$. Observe $H(a, W_j)$ has at most two elements satisfying equation (15), one or both of:

$$y_{2,1}(a) = x_{i_2} - \sqrt{r_i^2 - (x_{i_1} - a)^2} \quad (16)$$

or,

$$y_{2,2}(a) = x_{i_2} + \sqrt{r_i^2 - (x_{i_1} - a)^2}, \quad (17)$$

whenever $r_i^2 \geq (x_{i_1} - a)^2$. Using property (8) of the dirac delta function, and substituting with $y_{2,1}(a)$ and $y_{2,2}(a)$ obtained in (16) and (17), (14) becomes:

$$h(a, P_{r_i}, x_i) = \int_{H(a, j)} \delta(f(a, y_2, P_{r_i}), x_i) dy_2,$$

takes the values

$$\begin{cases} \int_{H(a, W_j)} \frac{\delta(y_2 - y_{2,1}(a))}{|f'(a, y_2, P_{r_i}, x_i)|} dy_2 \\ \text{if } y_{2,1}(a) \in H(a, W_j) \text{ but } y_{2,2}(a) \notin H(a, W_j) \\ \int_{H(a, W_j)} \frac{\delta(y - y_{2,2}(a))}{|f'(a, y_2, P_{r_i}, x_i)|} dy_2 \\ \text{if } y_{2,2}(a) \in H(a, W_j) \text{ but } y_{2,1}(a) \notin H(a, W_j) \\ \int_{H(a, W_j)} \frac{\delta(y_2 - y_{2,1}(a))}{|f'(a, y_2, P_{r_i}, x_i)|} + \frac{\delta(y_2 - y_{2,2}(a))}{|f'(a, y_2, P_{r_i}, x_i)|} dy_2 \\ \text{if both } y_{2,1}(a) \text{ and } y_{2,2}(a) \in H(a, W_j) \end{cases}$$

Define $I_1(a, W_j)$, the indicator function satisfying

$$I_1(a, W_j) = \begin{cases} 1 & \text{if } y_{2,1}(a) \in H(a, W_j) \\ 0 & \text{otherwise} \end{cases}$$

Similarly define $I_2(a, W_j)$, the indicator function satisfying

$$I_2(a, W_j) = \begin{cases} 1 & \text{if } y_{2,2}(a) \in H(a, W_j) \\ 0 & \text{otherwise} \end{cases}$$

Then, (14) becomes

$$h(a, P_{r_i}, x_i) = \frac{1}{|f'(a, y_{2,1}(a), P_{r_i}, x_i)|} I_1(a, W_j) + \frac{1}{|f'(a, y_{2,2}(a), P_{r_i}, x_i)|} I_2(a, W_j).$$

By substituting from (16), we get

$$\begin{aligned} & \frac{1}{|f'(a, y_2(a), P_{r_i}, x_i)|} \\ &= \frac{d_0 + ((x_{i_1} - a)^2 + r_i^2 - (x_{i_1} - a)^2)^{\frac{\beta}{2}}}{\beta \sqrt{r_i^2 - (x_{i_1} - a)^2} ((x_{i_1} - a)^2 + r_i^2 - (x_{i_1} - a)^2)^{\frac{\beta}{2} - 1}} \\ &= \frac{d_0 + r_i^\beta}{\beta \sqrt{r_i^2 - (x_{i_1} - a)^2} r_i^{\beta - 2}} = \frac{d_0 + r_i^\beta}{\beta r_i^{\beta - 2}} \cdot \frac{1}{\sqrt{r_i^2 - (x_{i_1} - a)^2}} \end{aligned}$$

Let $C_j = \{x \in \mathbb{R} \text{ such that } (x, y_{2,1}(x)) \in W_j\}$ and $C'_j = \{x \in \mathbb{R} \text{ such that } (x, y_{2,2}(x)) \in W_j\}$. Note that

$$x \in C_j \Rightarrow I_1(x, W_j) = 1 \text{ and } x \in C'_j \Rightarrow I_2(x, W_j) = 1.$$

Then,

$$\begin{aligned} p(\ln P_{r_i} | y \in W_j) P(y \in W_j) &= \frac{1}{A} \int_{C_j \cup C'_j} h(y_1, P_{r_i}, x_i) dy_1 \\ &= \frac{1}{A} \left(\int_{C_j} h(y_1, P_{r_i}, x_i) dy_1 + \int_{C'_j} h(y_1, P_{r_i}, x_i) dy_1 \right) \\ &= \frac{1}{A} \int_{C_j} \frac{1}{|f'(y_1, y_{2,1}(y_1), P_{r_i}, x_i)|} dy_1 \\ &+ \frac{1}{A} \int_{C'_j} \frac{1}{|f'(y_1, y_{2,2}(y_1), P_{r_i}, x_i)|} dy_1. \end{aligned} \quad (18)$$

Write

$$C_j = \bigcup_{\alpha=1}^s A_\alpha, \text{ with } \bigcap_{\alpha} A_\alpha = \emptyset, \text{ and } A_\alpha = [a_{1_\alpha}, a_{2_\alpha}], \quad (19)$$

$$C'_j = \bigcup_{\alpha=1}^{s'} A'_\alpha, \text{ with } \bigcap_{\alpha} A'_\alpha = \emptyset, \text{ and } A'_\alpha = [a'_{1_\alpha}, a'_{2_\alpha}]. \quad (20)$$

Equation (18) can then be written as the sum

$$\begin{aligned} & A \cdot p(\ln P_{r_i} | y \in W_j) P(y \in W_j) = \\ & \sum_{\alpha=1}^s \int_{A_\alpha} \frac{1}{|f'(y_1, y_{2,1}(y_1), P_{r_i}, x_i)|} dy_1 \\ & + \sum_{\alpha=1}^{s'} \int_{A'_\alpha} \frac{1}{|f'(y_1, y_{2,2}(y_1), P_{r_i}, x_i)|} dy_1 \\ &= \sum_{\alpha=1}^s \int_{a_{2_\alpha}}^{a_{1_\alpha}} \frac{1}{|f'(y_1, y_{2,1}(y_1), P_{r_i}, x_i)|} dy_1 \\ & + \sum_{\alpha=1}^{s'} \int_{a'_{1_\alpha}}^{a'_{2_\alpha}} \frac{1}{|f'(y_1, y_{2,2}(y_1), P_{r_i}, x_i)|} dy_1, \\ &= \sum_{\alpha=1}^s \int_{a_{2_\alpha}}^{a_{1_\alpha}} \frac{d_0 + r_i^\beta}{\beta r_i^{\beta - 2}} \cdot \frac{1}{\sqrt{r_i^2 - (x_{i_1} - y_1)^2}} dy_1 \\ & + \sum_{\alpha=1}^{s'} \int_{a'_{2_\alpha}}^{a'_{1_\alpha}} \frac{d_0 + r_i^\beta}{\beta r_i^{\beta - 2}} \cdot \frac{1}{\sqrt{r_i^2 - (x_{i_1} - y_1)^2}} dy_1 \\ &= \sum_{\alpha=1}^s \frac{d_0 + r_i^\beta}{\beta r_i^{\beta - 2}} \cdot \arctan \frac{x_{i_1} - a}{\sqrt{r_i^2 - (x_{i_1} - y_1)^2}} \Big|_{a_{2_\alpha}}^{a_{1_\alpha}} \\ & + \sum_{\alpha=1}^{s'} \frac{d_0 + r_i^\beta}{\beta r_i^{\beta - 2}} \cdot \arctan \frac{x_{i_1} - y_1}{\sqrt{r_i^2 - (x_{i_1} - y_1)^2}} \Big|_{a'_{1_\alpha}}^{a'_{2_\alpha}} \end{aligned}$$

Note that [10]

$$\begin{aligned} & \arctan \frac{x_{i_1} - a}{\sqrt{r_i^2 - (x_{i_1} - a)^2}} \\ &= \arctan \frac{x_{i_1} - a}{\sqrt{(x_{i_1} - a)^2 + (x_{i_2} - y_2(a))^2 - (x_{i_1} - a)^2}} \\ &= \arctan \frac{x_{i_1} - a}{x_{i_2} - y_2(a)} = \frac{\pi}{2} - \arctan \frac{x_{i_2} - y_2(a)}{x_{i_1} - a}. \end{aligned}$$

The conditional probability density becomes after simplifications,

$$\begin{aligned} p(\ln P_{r_i} | y \in W_j) P(y \in W_j) &= \frac{d_0 + r_i^\beta}{A \beta r_i^{\beta - 2}} \cdot \sum_{\alpha=1}^s \\ & \left(\frac{\pi}{2} - \arctan \frac{x_{i_2} - y_2(a_{2_\alpha})}{x_{i_1} - a_{2_\alpha}} - \frac{\pi}{2} + \arctan \frac{x_{i_2} - y_2(a_{1_\alpha})}{x_{i_1} - a_{1_\alpha}} \right) \\ & + \frac{d_0 + r_i^\beta}{A \beta r_i^{\beta - 2}} \sum_{\alpha=1}^{s'} \left(\frac{\pi}{2} - \arctan \frac{x_{i_2} - y_2(a'_{2_\alpha})}{x_{i_1} - a'_{2_\alpha}} - \frac{\pi}{2} \right. \\ & \left. + \arctan \frac{x_{i_2} - y_2(a'_{1_\alpha})}{x_{i_1} - a'_{1_\alpha}} \right) = \frac{d_0 + r_i^\beta}{A \beta r_i^{\beta - 2}} \left(\sum_{\alpha=1}^s \theta_\alpha + \sum_{\alpha=1}^{s'} \theta'_\alpha \right), \end{aligned}$$

where θ_α and θ'_α are the angles of the arcs in $S(W_j, r_i, x_i)$ described on distinct supports as in (19) and (20) when applicable. \blacksquare