On the Hybrid TOA/RSS Range Estimation in Wireless Sensor Networks

Angelo Coluccia, *Senior Member, IEEE* , and Alessio Fascista

***Abstract*— Distance estimation, which arises in many applications and especially in range-based localization, is addressed for joint received signal strength (RSS) and time of arrival (TOA) data. A statistical characterization of the joint maximum likelihood estimator, which is unavailable in closed-form, is provided together with a full performance assessment in terms of the actual mean squared error (MSE), in order to establish when hybrid estimation is superior compared to RSS-only or TOA-only estimation. Furthermore, a novel closed-form estimator is proposed based on an ad-hoc relaxation of the likelihood function, which removes the need to adopt iterative methods for hybrid TOA/RSS ranging and strikes a better bias-variance tradeoff for improved performance. A thorough theoretical analysis, corroborated by numerical simulations, shows the effectiveness of the proposed approach, which outperforms state-of-the-art solutions.**

***Index Terms*— Ranging, received signal strength, time of arrival, maximum likelihood, range-based localization.**

I. Introduction

DISTANCE estimation or *ranging* arises in many applications, in particular as preliminary step for range-based localization [1], [2]. Besides satellite-based positioning (GPS), range-based solutions are today popular in terrestrial systems [3]–[5], and in wireless sensor networks (WSNs) are typically preferred to range-free techniques [6], [7] for complexity reasons. Different information can be exploited, namely received signal strength (RSS), time (difference) of arrival (TOA/TDOA), angle of arrival (AOA) [8]–[10]. The latter is infrequent in WSNs since antenna arrays increase complexity and cost; RSS is very popular due to its ready availability [11]–[15].

Hybrid TOA/RSS approaches have been investigated in wireless [16]–[23] but also visible light [24] systems. Most works have focused on the *position* estimation problem, whose joint TOA/RSS Cramér-Rao lower bound (CRLB) is discussed in [25]. Less attention has received the problem of hybrid TOA/RSS *range* estimation. In [26] a fusion algorithm is proposed for range-based tracking using two independent processing chains (with moving average and Kalman filters) for RSS and TOA. By studying the uncertainty propagation, therein it is pointed out that RSS is more accurate at short ranges, while on the contrary TOA is more accurate at long ranges. The same conclusion is reached in [27] by analyzing the CRLB for joint TOA/RSS ranging with uncorrelated measurements. Although a more involved computation is needed when RSS and TOA are extracted from the same signal [28], [29], there is experimental evidence that the correlation is weak in practice [26], [30]; furthermore, in WSNs contexts multiple packets can be easily obtained, hence RSS and TOA can be separately measured, gaining full independence. Despite the resulting simplification of the joint TOA/RSS likelihood function, the Maximum Likelihood (ML) estimator does not admit an explicit form [30]. In [27] it is suggested to use RSS-only or TOA-only ranging respectively below and above a *critical distance*, giving no solution for all intermediate ranges. The state-of-the-art solution is built from Newton-Raphson iterations, which is shown to provide good performance [30]; however, results are demonstrated only in particular scenarios, and no theoretical tool is provided to establish whether hybrid ranging is convenient in a generic setting.

In this work we provide two kinds of contributions. In the first part of the paper, corresponding to Secs. II-III, we throughly analyze the problem from a theoretical perspective, complementing the CRLB-based analysis available in the literature [27]–[29] with a precise assessment of the actually-achievable performance in terms of mean squared error (MSE). In particular:

- a more precise MSE-based formula for the critical distance is derived, which (asymptotically) encompasses the CRLB-based one [27];
- the conditions under which hybrid estimation is advantageous compared to RSS-only or TOA-only estimation are established, which also explain why the joint ML estimator empirically performs comparably to TOA-only in [30] but not in general;
- a formal statistical analysis of the joint ML estimator, corroborated by simulations, is provided and its performance and issues discussed.

This analysis, which is completely general, demonstrates that joint TOA/RSS ranging is beneficial in many cases. In the second part of the paper, corresponding to Secs. IV-V, we provide the following contributions:

- a novel hybrid TOA/RSS estimation approach based on a relaxation of the likelihood function is proposed, which results in a closed-form estimator that avoids the drawbacks of iterative methods and provides a better bias-variance tradeoff;
- the structure of this estimator is further simplified for a convenient practical adoption in WSNs;

Manuscript received April 24, 2017; revised September 4, 2017; accepted October 20, 2017. Date of publication October 31, 2017; date of current version January 8, 2018. The associate editor coordinating the review of this paper and approving it for publication was K. Haneda. *(Corresponding author: Angelo Coluccia.)*

The authors are with the University of Salento, 73100 Lecce, Italy (e-mail: angelo.coluccia@unisalento.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2017.2766628

- a formal statistical analysis of the proposed estimator, corroborated by simulations in realistic conditions, is given which shows it outperforms the state-of-the-art competitor.

We report our conclusions in Sec. VI.

II. Hybrid TOA/RSS Range Estimation

We consider a network of sensors over a given area. Without loss of generality, we focus on a node pair, one transmitting packets to the other. Based on the received packets, the goal is to estimate the unknown distance d.

A. RSS-Based Ranging

The most general-purpose model for the average received power r_i in dB is the *path loss model* [31]:

$$r_i = P_0 - 10\alpha \log_{10} \frac{d}{d_0} + n_i \tag{1}$$

with α the *path loss exponent* and P_0 the received power at distance d_0; we can assume $d_0 = 1$ meter without loss of generality. The normally-distributed (in dB) random variables n_i, having zero-mean and variance σ_n^2, accounts for *shadowing* effects (log-normally distributed in natural units).[1] In general, if K independent measurements $r_1, \ldots, r_K$ are available, it is easy to show that a *sufficient statistic* to estimate d is the sample mean $r = \frac{1}{K}\sum_{i=1}^{K} r_i$, which lowers the variance to σ_n^2/K. To be general, we denote by σ_r^2 the final, irreducible level of uncertainty in r, and thus consider the following model

$$r \sim f_r(r; d) = \mathcal{N}(P_0 - 10\alpha \log_{10} d, \sigma_r^2) \tag{2}$$

where f_r denotes the probability density function (pdf) of r and $\mathcal{N}(m, s^2)$ is the Gaussian pdf with mean m and variance s^2. The model parameters vary with the environment; they can be estimated offline or by in-network "calibration" [14], [32]–[34]. Since our focus is on range estimation, we treat such parameters as given (deterministic) values; possible residual errors on the parameters will be taken into account by a greater value of σ_r^2. The ML estimator $\arg\max_d f_r$, very popular in range-based localization, is given by [14], [26]

$$\hat{d}_r^{\mathrm{ML}} \stackrel{\mathrm{def}}{=} 10^{\frac{P_0 - r}{10\alpha}} = \mathrm{e}^{\frac{P_0 - r}{a}} \tag{3}$$

where the more convenient parametrization $a \stackrel{\mathrm{def}}{=} \frac{10\alpha}{\log 10}$ is also introduced and hereafter adopted instead of α.

B. TOA-Based Ranging

If the departure time T_0 is known, i.e., a timestamp is attached to the packet just before it leaves the transmitter, the time of fly can be estimated up to the synchronization error. The measured time is thus

$$t_i = T_0 + \frac{d}{c} + w_i \tag{4}$$

where w_i is a zero-mean random term with variance σ_w^2. According to its features, a receiver has a different precision [26] reflected by σ_w^2. A Gaussian distribution is often adopted to model w_i, which typically includes measurement noise and calibration/synchronization errors.[2] Analogously to RSS, if K independent TOA measurements t_i are available, a sufficient statistic is given by the sample mean t; its final, irreducible level of uncertainty (variance) is likewise denoted by σ_t^2, thus

$$t \sim f_t(t; d) = \mathcal{N}\left(T_0 + \frac{d}{c}, \sigma_t^2\right). \tag{5}$$

Eq. (5) is widely-adopted in applications and has been validated on real data [21]–[23], [26], [30]. It can also cope with NLOS conditions by considering an additional offset in the mean, to be estimated and compensated by "calibration" [30]. Again, a greater σ_t^2 will be considered to take into account residual calibration errors.

The most natural estimator of d is clearly

$$\hat{d}_t \stackrel{\mathrm{def}}{=} c\,(t - T_0) \equiv \hat{d}_t^{\mathrm{ML}} \tag{6}$$

which, since (5) is a linear Gaussian model, is also the ML estimator irrespective of the value of σ_t [26].[3]

C. Joint Maximum Likelihood Estimator

The hybrid TOA/RSS ML estimator is obtained by maximizing the joint pdf of the RSS and TOA observations. Using (2) and (5) the likelihood function is

$$\begin{aligned}\Lambda(r, t|d) &= f_r(r|d)\, f_t(t|d) \\ &= \frac{1}{2\pi\,\sigma_r \sigma_t} \mathrm{e}^{-\frac{1}{2\sigma_t^2\sigma_r^2}\left[(r - P_0 + a\log d)^2\sigma_t^2 + \left(t - T_0 - \frac{d}{c}\right)^2 \sigma_r^2\right]}\end{aligned} \tag{7}$$

Eq. (7) is the exact likelihood if r and t are measured from independent sources, namely different packets, which can be easily obtained in WSNs. Furthermore, as mentioned in Sec. I, experiments show that RSS and TOA extracted from the same signal are anyway weakly correlated [26], [30], hence the assumption of independence is not detrimental to generality. Adopting the notation $\boldsymbol{y} = [r \;\; t]^T$, the problem can be written as

$$\hat{d}_{\mathrm{JML}} = \arg\max_d \Lambda(\boldsymbol{y}|d) \equiv \arg\min_d L(\boldsymbol{y}; d)$$

where L is equivalent to the negative log-likelihood, i.e.,

$$L(\boldsymbol{y}; d) = (r - P_0 + a\log d)^2\, \sigma_t^2 + \left(t - T_0 - \frac{d}{c}\right)^2 \sigma_r^2. \tag{8}$$

The joint TOA/RSS ML range estimator is the function $h(\boldsymbol{y})$ satisfying the first-order optimality condition

$$\left.\frac{\partial L(\boldsymbol{y}; d)}{\partial d}\right|_{d = h(\boldsymbol{y})} = 0 \tag{9}$$

[1] Since fast signal variations due to small-scale fading are mitigated by time averaging in r_i ([11], [32], [34], see also [35] for details about RSS computation) model (1) is indeed sufficient to represent the mean received power as affected by the sole shadowing component (n_i), as confirmed by extensive measurements [31], [36]–[38].

[2] A popular lower bound for the variance is $\frac{1}{8\pi^2 \beta\, \mathrm{SNR}}$, where β is a normalized mean-squared measure of bandwidth [27], [39]. Thus, other things being equal, the receiver will estimate the time of fly more accurately as SNR and/or bandwidth increase.

[3] The TOA model can be equivalently rewritten in terms of the estimated and true distances, i.e., $\hat{d}_t = d + w$, as e.g. in [19].

i.e., $\hat{d}_{\text{JML}} = h(\boldsymbol{y})$. This leads to the equation

$$\frac{\sigma_t^2}{d}(r - P_0 + a\log d)a - \frac{\sigma_r^2}{c}\left(t - T_0 - \frac{d}{c}\right) = 0 \quad (10)$$

which however does not admit an explicit solution in d.

III. Analysis of RSS-Only, TOA-Only, and Joint ML Estimators

A. Biased vs. Unbiased RSS-Based Ranging

While the ML procedure guarantees asymptotic unbiasedness and normality [14] for both TOA-based and RSS-based estimators, an important difference is that the former is always unbiased ($\mathrm{E}[\hat{d}_t] = d$) while the latter is only asymptotically unbiased — which for the problem at hand means $\sigma_r \to 0$. The bias amounts to $\mathrm{E}[\hat{d}_r^{\text{ML}} - d] = \left(\mathrm{e}^{\frac{\sigma_r^2}{2a^2}} - 1\right)d$; thus, unbiasedness can be obtained through the modified estimator [32], [40], [41]

$$\hat{d}_r \stackrel{\text{def}}{=} 10^{\frac{1}{a\log 10}\left(P_0 - r - \frac{\sigma_r^2}{a}\right)} = \mathrm{e}^{\frac{P_0 - r}{a} - \frac{\sigma_r^2}{a^2}}. \quad (11)$$

Statistical bias is an issue in ranging [33], [42], [43]; at the same time, the well-known *bias-variance tradeoff* indicates that a biased estimator exhibits a lower variance compared to an unbiased one, which may or may not ultimately determine better performance in terms of mean squared estimation error (MSE).[4] We were not able to find a reference where it has been clarified if $\hat{d}_r$ is ultimately more accurate than $\hat{d}_r^{\text{ML}}$, thus before continuing we provide such an assessment. For $\hat{d}_r^{\text{ML}}$ we have

$$\mathrm{MSE}[\hat{d}_r^{\text{ML}}] = \mathrm{E}\left[\left(\hat{d}_r^{\text{ML}} - d\right)^2\right] = \left(\mathrm{e}^{\frac{2\sigma_r^2}{a^2}} - 2\mathrm{e}^{\frac{\sigma_r^2}{2a^2}} + 1\right)d^2 \quad (12)$$

and analogous computations yield the MSE of $\hat{d}_r$:

$$\mathrm{MSE}[\hat{d}_r] = \mathrm{VAR}[\hat{d}_r] = \left(\mathrm{e}^{\frac{\sigma_r^2}{a^2}} - 1\right)d^2. \quad (13)$$

Asking whether $\mathrm{E}[(\hat{d}_r^{\text{ML}} - d)^2] \stackrel{?}{>} \mathrm{E}[(\hat{d}_r - d)^2]$ means considering the inequality $\mathrm{e}^{\frac{2\sigma_r^2}{a^2}} - \mathrm{e}^{\frac{\sigma_r^2}{a^2}} - 2\mathrm{e}^{\frac{\sigma_r^2}{2a^2}} + 2 > 0$; since this is always satisfied, we conclude that *the unbiased $\hat{d}_r$ dominates the biased $\hat{d}_r^{ML}$* in terms of MSE performance, hence the former should be always preferred to the latter.

B. Critical Distance Analysis

Let us now consider the Cramér-Rao lower bound (CRLB) for the different estimators. From (2) (or [27])

$$\mathrm{CRLB}_{\text{RSS}}(d) = -\frac{1}{\mathrm{E}\left[\frac{\partial}{\partial d^2}\log f_r\right]} = \frac{\sigma_r^2}{a^2}d^2. \quad (14)$$

Comparing (13) and (14) gives $\frac{\sigma_r^2}{a^2} \geq \log\left(1 + \frac{\sigma_r^2}{a^2}\right)$; since $\log(1+x) < x, \forall x > -1 \wedge x \neq 0$, we conclude that $\mathrm{MSE}[\hat{d}_r] > \mathrm{CRLB}_{\text{RSS}}(d)$. This is a point of difference with respect to $\hat{d}_t$, for which (5) (or [27]) indicates that

$$\mathrm{CRLB}_{\text{TOA}}(d) = -\frac{1}{\mathrm{E}\left[\frac{\partial}{\partial d^2}\log f_t\right]} = (c\sigma_t)^2 \quad (15)$$

and at the same time

$$\mathrm{MSE}[\hat{d}_t] = \mathrm{E}\left[\left(\hat{d}_t - d\right)^2\right] = \mathrm{VAR}[\hat{d}_t] = (c\sigma_t)^2 \quad (16)$$

thus TOA-alone ranging is CRLB-*efficient* for any σ_t.

The CRLB for the joint TOA/RSS case (ref. (7)) is

$$\mathrm{CRLB}_{\text{JOINT}}(d) = \frac{\sigma_r^2(c\sigma_t)^2 d^2}{d^2\sigma_r^2 + a^2(c\sigma_t)^2} = \frac{1}{\frac{1}{\mathrm{CRLB}_{\text{RSS}}(d)} + \frac{1}{\mathrm{CRLB}_{\text{TOA}}(d)}}. \quad (17)$$

A simple calculation shows that for small d the McLaurin's expansion of (17) returns (14) plus higher order terms (h.o.t.); also, the limit for large d is (15). It is intuitive that if one measurement is much more accurate than the other one, then there is little advantage in hybrid ranging, as observed in [26], [27]. However, a general analysis is missing in the literature, to precisely establish whether hybrid ranging can yield superior performance in a given operational environment. Indeed, recall that the values of $c\sigma_t$ and σ_r depend on physical-layer aspects such as SNR, bandwidth, and synchronization quality, as well as on the number K of independent measurements available, as discussed in Secs. II-A and II-B. Fig. 1 reports a graph for $c\sigma_t = \sigma_r = 6$, a noticeable case with same accuracy for the two measurements: it is apparent that the joint CRLB (17) is significantly below the TOA-only (15) and RSS-only (14) CRLBs, indicating a margin to obtain a better estimate from the joint statistic (t, r).

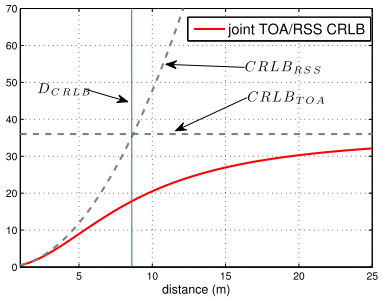


Fig. 1. Comparison between the RSS-only, TOA-only, and joint TOA/RSS CRLB, with associated critical distance, $c\sigma_t = \sigma_r = 6$.

The figure also shows the value of the *critical distance* $D_{\text{CRLB}} = \frac{a}{\gamma}$ obtained in [27] by equating (14)-(15), where we denoted by $\gamma \stackrel{\text{def}}{=} \frac{\sigma_r}{c\sigma_t}$ the ratio between RSS and TOA standard deviations. Actually, a more precise definition is allowed by the MSEs (13)-(16) we derived above, i.e.,

$$D_{\text{MSE}} = \frac{c\sigma_t}{\sqrt{\mathrm{e}^{\frac{\sigma_r^2}{a^2}} - 1}}. \quad (18)$$

For small $\frac{\sigma_r}{a}$ clearly (18) coincides with D_{CRLB}, however in general $D_{\text{MSE}} < D_{\text{CRLB}}$ with deviations up to 20%. Hereafter we

[4] This tradeoff has a major role in most statistical problems (see e.g. [44]) and is nowadays regarded as a key element in connecting the classical Fisher's "unbiasedness culture", in particular ML and CRLB, to different estimation approaches able to perform well with small sample size (more generally, under unfavorable statistical conditioning, see the interesting discussion in [45]). We will see that our proposed estimator indeed leverages bias to achieve better MSE.

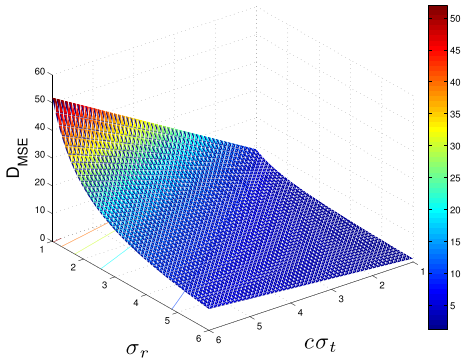


Fig. 2. Values of the critical distance D_{MSE} (in meters) as function of $c\sigma_t$ and σ_r, for $\alpha = 2$.

stick to D_{MSE} as definition of critical distance, since it better reflects the actual performance.[5]

Fig. 2 shows that D_{MSE} may significantly vary depending on $c\sigma_t$ and σ_r. For d much smaller [greater] than D_{MSE} the joint estimation is practically equivalent to RSS-only [TOA-only] estimation: thus, if σ_r and $c\sigma_t$ are very different, D_{MSE} accordingly turns very small or very large hence $d < D_{\text{MSE}}$ or $d > D_{\text{MSE}}$ for practically any d. On the other hand, the typical inter-node separation in the WSN context, from a few meters up to a few tens of meters, is well within the region where joint TOA/RSS estimation can be beneficial. Similar considerations hold true for different values of α (equivalently, a).

The analysis above explains also the empirical findings in [30], where the accuracy of the joint estimator turns out to be comparable to TOA-only. For the scenario considered therein, the equivalent figures are $c\sigma_t \approx 5.1$ and $\sigma_r \approx 6.8$ for an indoor environment, where α is typically small; this leads to D_{MSE} of just a few meters, while the reported results consider as minimum distance of interest a value of 5 meters. Thus, for such a span (higher than the critical distance) the RSS is not much informative, and joint estimation is not better than TOA-only; as discussed above, however, there are many other cases of interest in which this does not hold true.

C. Analysis of the Joint TOA/RSS ML Estimator

As reviewed in Sec. II-C, $\hat{d}_{\text{JML}}$ does not admit a closed-form, thus no exact formal analysis is possible. We give a characterization based on a perturbation analysis.

Proposition 1: The bias of $\hat{d}_{JML}$, up to h.o.t., is given by

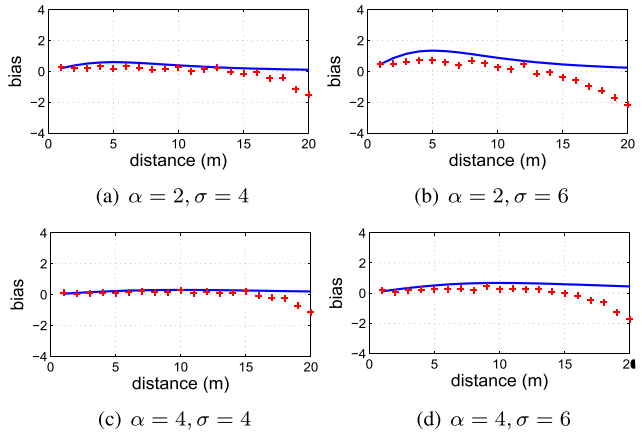
$$B_{JML} = (c\sigma_t)^2 \left[a\left(a+\frac{1}{2}\right)(c\sigma_t)^2\sigma_r^2 - \frac{4a}{c^2}\right]\frac{d}{[x(d)]^2} \approx \frac{[\text{CRLB}_{TOA}(d)]^2\,\text{CRLB}_{RSS}(d)}{[\text{CRLB}_{TOA}(d)+\text{CRLB}_{RSS}(d)]^2}d \quad (19)$$

where $x(d) = a^2(c\sigma_t)^2 + \sigma_r^2 d^2$.

Proof: see Appendix A. □

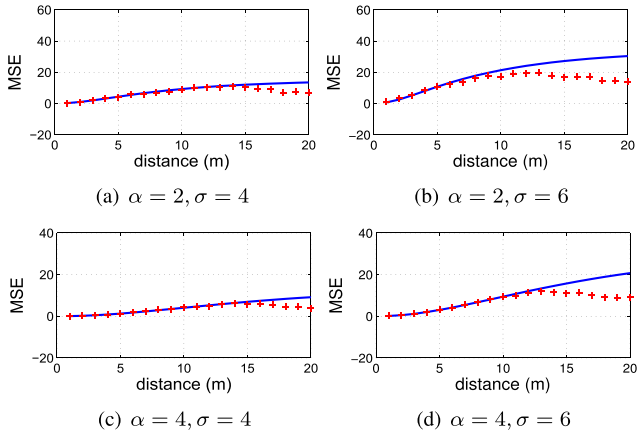
Corollary $\sigma_r = c\sigma_t \overset{def}{=} \sigma$: $B_{\text{JML}} \approx \frac{a^2 d}{(a^2+d^2)^2}\sigma^2$. ◇

[5]In fact, D_{CRLB} corresponds to the asymptotic performance of an (ideal) estimator, while D_{MSE} is related to the actual MSE of specific computable estimators, with the (unbiased) estimator $\hat{d}_r$ never attaining the corresponding CRLB.



(a) $\alpha = 2, \sigma = 4$ (b) $\alpha = 2, \sigma = 6$ (c) $\alpha = 4, \sigma = 4$ (d) $\alpha = 4, \sigma = 6$

Fig. 3. Approximate formulas for the bias of the joint ML estimator $\hat{d}_{\text{JML}}$, in comparison to Monte Carlo simulation points.



(a) $\alpha = 2, \sigma = 4$ (b) $\alpha = 2, \sigma = 6$ (c) $\alpha = 4, \sigma = 4$ (d) $\alpha = 4, \sigma = 6$

Fig. 4. Approximate formulas for the MSE of the joint ML estimator $\hat{d}_{\text{JML}}$, in comparison to Monte Carlo simulation points.

Fig. 3 reports B_{JML} for different values of the parameters: it appears in good agreement with the points obtained through Monte Carlo simulations (10000 trials), with some deviations for large distances which however amount to a few percent in relative terms. This is quite remarkable, considering that $\hat{d}_{\text{JML}}$ has not a closed-form to be analyzed. From the results, we observe that in most cases the bias is not severe compared to distances of the order of meters. Looking at (19), it is clear that as σ_r and/or $c\sigma_t$ decrease, the corresponding CRLBs become smaller and eventually the bias becomes negligible.

Proposition 2: The MSE of $\hat{d}_{JML}$, up to h.o.t., is given by

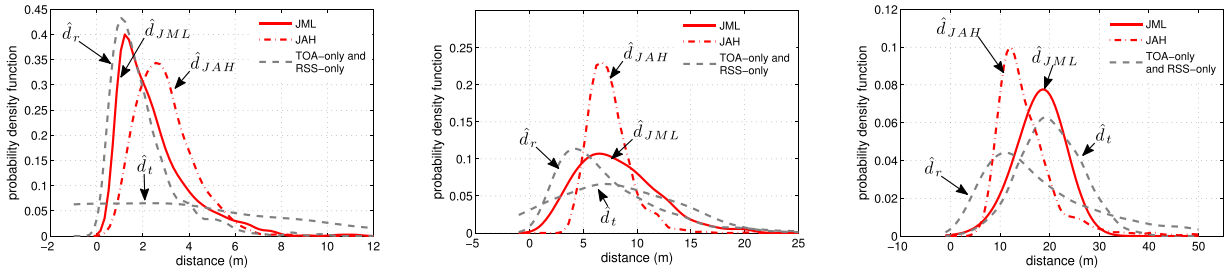
$$\text{MSE}[\hat{d}_{JML}] = (c\sigma_t)^2\sigma_r^2\frac{d^2}{x(d)} + B_{JML}^2 = \frac{\text{CRLB}_{TOA}(d)\,\text{CRLB}_{RSS}(d)}{\text{CRLB}_{TOA}(d)+\text{CRLB}_{RSS}(d)} + B_{JML}^2 \quad (20)$$

where $x(d)$ and B_{JML} are given in Proposition 1.

Proof: see Appendix B. □

Corollary $\sigma_r = c\sigma_t \overset{def}{=} \sigma$: $\text{MSE}[\hat{d}_{\text{JML}}] \approx \left[\frac{\sigma^2}{a^2+d^2} + \frac{a^4\sigma^4}{(a^2+d^2)^4}\right]d^2$. ◇

Fig. 4 shows that the accuracy of the formulas is very good, again with some relatively modest deviations for large d. Notice that the first addend in the expressions given in (20)



(a) $d = 2$ meters ($< D_{\text{MSE}}$) (b) $d = D_{\text{MSE}} \approx 7.7$ meters (c) $d = 20$ meters ($> D_{\text{MSE}}$)

Fig. 5. Pdfs of the considered range estimators for $\sigma_r = c\sigma_t = 6$.

is related to the variance of $\hat{d}_{\text{JML}}$, and it coincides (up to h.o.t.) with the $\text{CRLB}_{\text{JOINT}}$ (17). However, the actual MSE is higher due to the presence of the additive term B^2_{JML} related to the bias. While this is of course fully consistent with the (only-)asymptotic CRLB-efficiency property of ML estimation, on the other hand the analysis performed above using Proposition 1 shows that $\hat{d}_{\text{JML}}$ is only slightly biased, i.e., it generally performs not far from the CRLB. It is important to remark that the latter is a theoretical benchmark for unbiased estimators, while the bias-variance tradeoff indicates that a biased estimator can be more accurate for finite (especially small) sample — see comment in Sec. III-A; following this idea, in the second part of this paper we devise a novel ad-hoc approach for hybrid TOA/RSS ranging, which can exhibit MSE even smaller than $\hat{d}_{\text{JML}}$ and admits a closed-form.

To qualitatively introduce this point, and at the same time wrap-up the outcomes of the analysis performed in this section, Fig. 5 shows the pdfs of TOA-only, RSS-only, joint ML, and proposed ad-hoc estimators — the latter denoted by $\hat{d}_{\text{JAH}}$, thoroughly discussed in the next section — for three different ranges; in particular, we consider a distance smaller, equal, and larger than the critical distance, in order to expose the different behaviors of the estimators. The pdf of the joint TOA/RSS ML estimator $\hat{d}_{\text{JML}}$, for which a closed-form is not available, is evaluated through normal kernel smoothing [46], and so does for the pdf of the proposed estimator.[6] From Fig. 5(a) it is evident that for $d < D_{\text{MSE}}$ the pdf of $\hat{d}_{\text{JML}}$ is close to the one of the RSS-only estimator $\hat{d}_r$, due to the fact that the information brought by the TOA in this case is not significant, as visible from the very dispersed pdf of the latter. *Vice versa*, in Fig. 5(c) the pdf of $\hat{d}_{\text{JML}}$ is close to the one of the TOA-only estimator $\hat{d}_t$, while the pdf of $\hat{d}_r$ is more dispersed. As predicted by the analysis, around the critical distance D_{MSE} the $\hat{d}_{\text{JML}}$ benefits from the information brought by both RSS and TOA, and in fact its pdf is a "merge" of the ones of $\hat{d}_r$ and $\hat{d}_t$. Interestingly, the pdf of the proposed ad hoc estimator $\hat{d}_{\text{JAH}}$ is more alike the pdf of $\hat{d}_{\text{JML}}$, but with an important difference: it trades-off an higher bias with a reduced variance (though this might be not immediately visible due to the skewness of the distribution), i.e., it is less dispersed than the $\hat{d}_{\text{JML}}$ in the whole range of considered distances. Proving that this ultimately determines an actual advantage in terms of MSE performance is part of the following section, which includes the derivation of the proposed ad-hoc estimator as well as its formal statistical analysis.

[6]It is conversely a simple matter to derive the pdfs of $\hat{d}_t$ and $\hat{d}_r$: the former is a Gaussian due to the linear relationship with t (ref. (5)); the latter is a nonlinear transformation of r that, by the change of variable theorem, is equivalent to a log-Normal distribution.

IV. Ad-Hoc Hybrid TOA/RSS Estimator

Computation of $\hat{d}_{\text{JML}}$ requires the iterative minimization of L (8). It is well-known that optimization routines are sensitive to the choice of the initial point, which may result in significantly slower or even missing convergence. Also, naive implementations (as the ones that can be reasonably performed by practitioners lacking expertise in numerical optimization techniques [47]) may have problems when deployed on WSN nodes; indeed, often the embedded processors are even fixed-point arithmetic. These drawbacks explain the appeal of closed-form solutions. To the best of our knowledge, the Newton-Raphson based solution [30] is the sole closed-form hybrid TOA/RSS estimator available in the literature. We propose below an alternative ad-hoc approach based on a relaxation of the quadratic term in the optimality condition (10).

A. Derivation of the Proposed Approach

To derive the proposed estimator, first rewrite (10) as

$$\log d = \underbrace{-\frac{\gamma^2}{a^2}d^2 + \frac{\gamma^2 c\,(t - T_0)}{a^2}d - \frac{r - P_0}{a}}_{z(d;\,\mathbf{y})}. \qquad (21)$$

Notice that in (21) the dependency is only on γ, not on σ_r and $c\sigma_t$ separately, and that the RSS-only and TOA-only estimators given by (3) and (6) are obtained for $\gamma \to 0$ and $\gamma \to \infty$, respectively. The idea is to approximate the quadratic right-hand side of this equation, denoted by $z(d;\,\mathbf{y})$, by relaxing it to an affine function of the form $md+q$. Clearly, there are infinite possible choices; we first provide a general result based on a set of n selected "control points" d_i, $i = 1, \ldots, n$, which are used to find the parameters m and q minimizing the sum of squared residuals from the set of $z_i \stackrel{\text{def}}{=} z(d_i;\,\mathbf{y})$, i.e., an ordinary least-squares approach similar to the one used in curve fitting and regression. The choice of such points is based on design considerations about the span of interest, namely it encodes the will to give more importance to some regions compared to

others — the denser the points, the better the approximation in that region. Solving the resulting optimization problem

$$(m^\star, q^\star) = \arg\min_{m,q} \sum_{i=1}^{n} (z_i - m d_i - q)^2 \tag{22}$$

leads to the following relaxed optimality condition:

$$\log d = m^\star d + q^\star. \tag{23}$$

The solution of (23), denoted by $\hat{d}_{\text{JAH}}$, can be interpreted as the intersection point between the curve $\log(d)$ and the line $m^\star d + q^\star$, whereas the optimal solution $\hat{d}_{\text{JML}}$ is the intersection point between $\log(d)$ and $z(d; \boldsymbol{y})$ (ref. eq. (21)). We have the following result.

Proposition 3: Assuming $c(t - T_0) < \delta = \frac{\delta_3 - \delta_1\delta_2}{\delta_2 - \delta_1^2}$, *where* $\delta_k = \frac{1}{n}\sum_{i=1}^{n} d_i^k$, $k = 1, 2, 3$, *the estimator* $\hat{d}_{JAH}$ *based on the relaxation* (23) *of* (21) *is given by*

$$\hat{d}_{JAH} = \frac{1}{\varphi} W_0\left(\varphi\, \mathrm{e}^{\psi}\right)$$
$$\varphi = \frac{\gamma^2}{a^2}\Big[c(T_0 - t) + \delta\Big], \quad \psi = \frac{\gamma^2}{a^2}\Big[\frac{a}{\gamma^2}(P_0 - r) + \tilde{\delta}\Big] \tag{24}$$

with $W_0(\cdot)$ *the principal branch of the Lambert W-function and* $\tilde{\delta} = \delta_2 - \delta_1\delta$.

Proof: see Appendix C. □

The values δ_k are the k-th power of the arithmetic, quadratic, and cubic generalized means, also called Hölder means. A special choice of $\{d_1, \ldots, d_n\}$ is the set of *uniformly* spaced points in an interval $[A, B]$, i.e.,

$$d_i \stackrel{\text{def}}{=} A + (i-1)\frac{B - A}{n - 1} \tag{25}$$

for $i = 1, \ldots, n$. In that case, δ and $\tilde{\delta}$ are as follows.

Proposition 4: If d_i*s are uniformly spaced in* $[A, B]$,

$$\delta = A + B, \quad \tilde{\delta} = \frac{A^2 + B^2}{2} - \frac{(B - A)^2}{3}\frac{n - \frac{1}{2}}{n - 1}$$

Proof: see Appendix D. □

Some remarks are now in order. Firstly notice that, while from (21) one might expect that the dependency on data variance is only through γ, the fact that φ and ψ are generally dependent on t and r alone, respectively, is quite remarkable. Secondly, once the set of control points $\{d_1, \ldots, d_n\}$ has been chosen, related quantities $\vec{d} \equiv \delta_1, \delta_2, \delta_3$, hence $\tilde{\delta}$ and δ can be pre-computed offline, being independent of $\boldsymbol{y}$. Proposition 4 further simplifies the estimator, since the dependency remains only on the edges of the interval and it is not necessary to compute the generalized means explicitly. This result gives also insights on the impact of the choice of n: interestingly, φ is independent of n, while ψ has a dependency through the term $\frac{n-\frac{1}{2}}{n-1}$. For n sufficiently large, which corresponds to a fine-grained sampling of the curve $z(d; \boldsymbol{y})$, such a dependency practically vanishes; in the limit, this is tantamount to taking a continuum of points, i.e., the special (continuous) version of problem (22)

$$(m^\star, q^\star) = \arg\min_{m,q} \int_A^B (z(x; \boldsymbol{y}) - mx - q)^2 \, \mathrm{d}x.$$

In fact, it can be shown that direct solution of the problem above returns the same result of Propositions 3-4 as $n \to \infty$ (proof omitted for the sake of conciseness). The additional nice feature of this special choice is that

$$\lim_{n\to\infty} \psi = \frac{P_0 - r}{a} - \frac{\gamma^2}{a^2}\frac{(A + B)^2 + 2AB}{6} \tag{26}$$

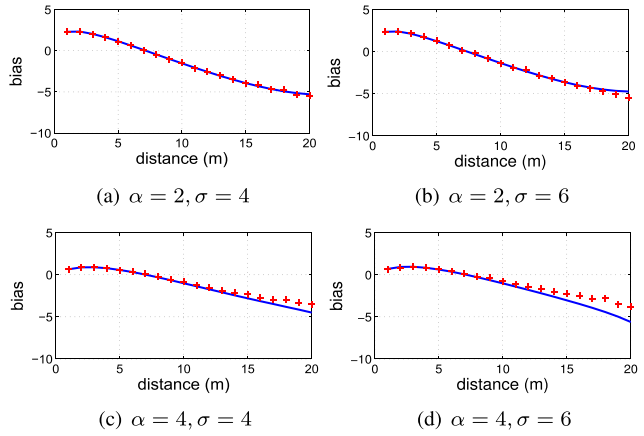
hence $\hat{d}_{\text{JAH}}$ can be simply computed with a single equation depending only on two design parameters, i.e., A, B. This is a remarkable property of the derived estimator; indeed, if at a first glance one may interpret A, B as the limits of the fitting interval, namely the range where the linear approximation better fits the quadratic form in (21), a deeper analysis reveals that $\hat{d}_{\text{JAH}}$ can be ultimately parametrized by A, B, which then act as two generic degrees of freedom that can be optimized for improved performance, as shown in Sec. V.

It is worth noticing that φ and ψ are affine transformations of Gaussian variables, hence are normally-distributed and easily characterizable in exact form. Thus, it is not difficult to understand the span of variation for the argument of the Lambert W-function in (24), as well as to establish when the hypothesis of Proposition 3 is satisfied. The term $c(t - T_0)$ actually coincides with the expression of $\hat{d}_t$, whose statistical characterization has been given in Sec. II-B. Thus, the probability for the condition in Proposition 3 to be fulfilled is $1 - Q\left(\frac{\delta - d}{c\sigma_t}\right)$, where $Q(\cdot)$ is the tail probability of the standard Gaussian with, more explicitly, $\delta = A + B$ for uniformly-spaced control points. Such a probability is typically very close to one, therefore to ease the presentation in the following we will consider only $\hat{d}_{\text{JAH}}$ given in Proposition 3. Though, the exact expression for the general case requires only minor modifications (see Appendix C).

As to the Lambert W-function, [48], [49] discuss how it can be computed efficiently. Even more simply, since $W_0(x)$ grows as slow as $\log x - \log\log x$, a look-up table (LUT) can be used to store its values only for integer arguments, without introducing appreciable differences in the results. Hereafter we will always refer to this very simple and immediate computation of the Lambert W-function, which requires only a few kilobytes of memory hence is particularly appealing for the low-cost WSN context where computational resources can be very limited. Notice that this is a prominent motivation in the literature on localization in WSNs, thus represents a major point of merit compared to the iterative computation of the joint ML estimator.

B. Analysis

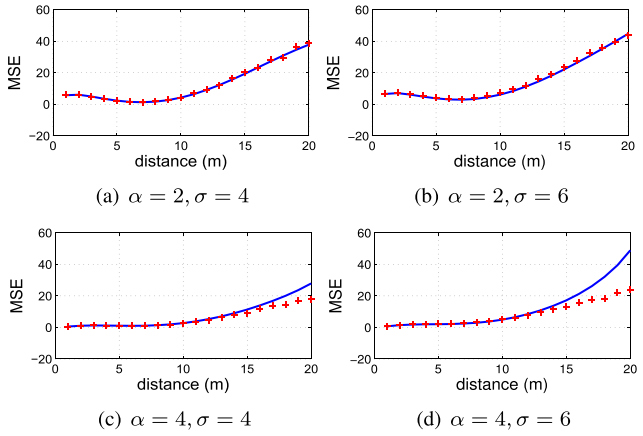
We have qualitatively seen in Fig. 5 that the proposed ad hoc estimator $\hat{d}_{\text{JAH}}$ shows two interesting features: (i) its pdf is more similar to the pdf of $\hat{d}_{\text{JML}}$ than of the individual TOA-only and RSS-only estimators, and (ii) it trades-off a higher bias with a reduced variance (though, as observed, this might be not immediately appreciable by visual inspection due to the skewness of the distribution), i.e., it is less dispersed than the $\hat{d}_{\text{JML}}$ in the whole range of considered distances. In this section we aim at establishing quantitatively if this ultimately determines an actual advantage in terms of MSE performance.



(a) $\alpha = 2, \sigma = 4$ (b) $\alpha = 2, \sigma = 6$

(c) $\alpha = 4, \sigma = 4$ (d) $\alpha = 4, \sigma = 6$

Fig. 6. Approximate formulas for the bias of the hybrid TOA/RSS ad-hoc estimator $\hat{d}_{\text{JAH}}$, in comparison to Monte Carlo simulation points.



(a) $\alpha = 2, \sigma = 4$ (b) $\alpha = 2, \sigma = 6$

(c) $\alpha = 4, \sigma = 4$ (d) $\alpha = 4, \sigma = 6$

Fig. 7. Approximate formulas for the MSE of the hybrid TOA/RSS estimator $\hat{d}_{\text{JAH}}$, in comparison to Monte Carlo simulation points.

To do so, we perform a formal analysis as the one we provided for the joint ML estimator, giving two results analogous to Propositions 1-2; as corollary, they can be trivially particularized to $\sigma_r = c\sigma_t$ by letting $\gamma = 1$.

Proposition 5: The bias of $\hat{d}_{JAH}$, up to h.o.t., is given by

$$B_{JAH} = \xi - d + \frac{\sigma_r^2}{2}\frac{\lambda^2}{a^2(\lambda+\mu)^3} + \frac{(c\sigma_t)^2}{2}\left(\frac{\gamma^2}{a^2\mu}\right)^2\left[2\xi - \frac{1}{\lambda+\mu}\left(1+\frac{2\mu\lambda+\mu^2}{(\lambda+\mu)^3}\right)\right]$$

where $\mu = \frac{\gamma^2}{a^2}(\delta - d)$, $\xi = \frac{1}{\mu}W_0\left(\mu\, d\, \mathrm{e}^{\frac{\gamma^2}{a^2}\tilde{\delta}}\right)$ *and* $\lambda = d\, \mathrm{e}^{\mu\xi - \frac{\gamma^2}{a^2}\tilde{\delta}}$.

Proof: see Appendix E. □

Fig. 6 reports the bias of the proposed ad-hoc estimator $\hat{d}_{\text{JAH}}$, where the predicted values are in very good agreement with the Monte Carlo points. Some deviations occur only in the cases (c)-(d) and limited to large distances. Looking at the results, the bias is more pronounced compared to the one of $\hat{d}_{\text{JML}}$ (Fig. 3). As mentioned, this may or may not correspond to a better balance in the bias-variance tradeoff, so to assess the ultimate performance one has to look at the MSE.

Proposition 6: The MSE of $\hat{d}_{JAH}$, up to h.o.t., is given by

$$\text{MSE}[\hat{d}_{JAH}] = \frac{\sigma_r^2}{a^2(\lambda+\varphi)^2} + \left(\frac{c\sigma_t\gamma^2}{a^2\varphi}\right)^2\left(\xi - \frac{1}{\lambda+\varphi}\right)^2 + B_{JAH}^2$$

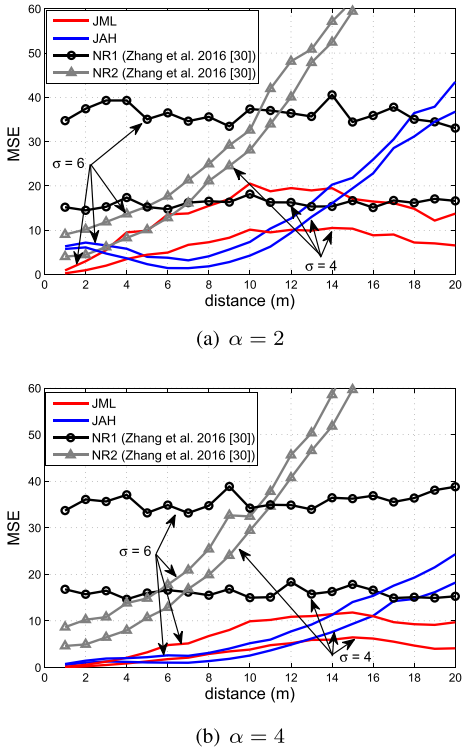
where ξ, λ, and B_{JAH} are given in Proposition 5.

Proof: see Appendix F. □

Fig. 7 shows that the predictions are very accurate, again with some deviations only in the cases (c)-(d) and limited to large distances. A direct comparison between the MSEs of $\hat{d}_{\text{JAH}}$ and $\hat{d}_{\text{JML}}$, as well as against the natural competitor [30], is performed below.

V. Performance Assessment

We first consider a neutral, general-purpose setting of A and B, i.e., for a maximum inter-node distance of 20 meters we set $A = -1$ and $B = 37$ which is tantamount to not optimizing at all the estimator for the specific span. We will see later that performance can further improve by a more specific tuning.



(a) $\alpha = 2$

(b) $\alpha = 4$

Fig. 8. Performance comparison (MSE) between the proposed ad-hoc estimator $\hat{d}_{\text{JAH}}$, the (iterative) ML benchmark $\hat{d}_{\text{JML}}$, and the two competitors proposed in [30], for $\sigma = 4$ and $\sigma = 6$.

For the assessment we consider, without loss of generality, the case $\sigma_r = c\sigma_t \stackrel{\text{def}}{=} \sigma$, the most interesting one since it has no *a priori* accuracy imbalance towards either RSS or TOA ($\gamma = 1$). Fig. 8 reports the MSE for the cases $\alpha = 2$ and $\alpha = 4$ and two values of σ ($\sigma = 4$ and $\sigma = 6$), as function of the inter-node distance d to be estimated. Notice that such values are representative of most application contexts of WSNs, varying from fairly open spaces, especially

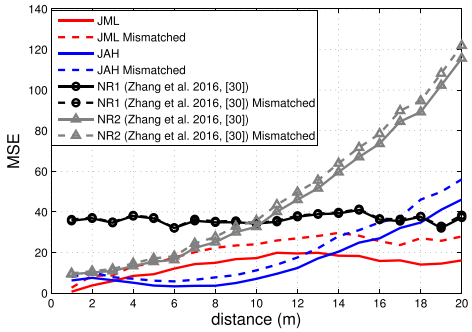
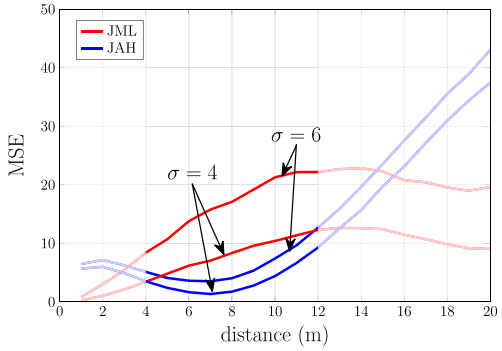


Fig. 9. Performance comparison under parameter mismatches.

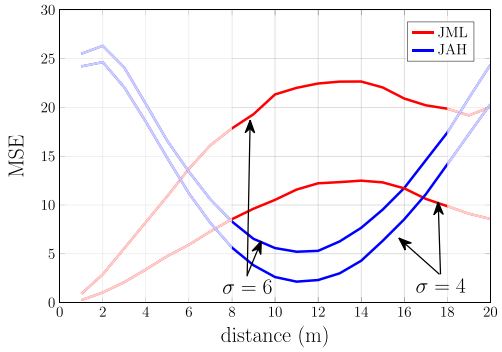
outdoor, to harsh cluttered environments full of metal and obstructions, including moving people, furniture, etc. For comparison, the state-of-the-art closed-form competitor [30] and the joint ML estimator $\hat{d}_{\text{JML}}$ obtained by numerical optimization are also reported; we remark that the latter should not be considered as a direct competitor, but only as a benchmark, since it is not a closed-form estimator. Results clearly show that the proposed ad-hoc estimator outperforms the competitor, and is even better than the ML benchmark in a significant part of the span. Results for other values of the parameters confirm the same findings.

We also analyzed the impact of possible errors in the assumed parameters. We perturbed the values of α, P_0, σ_r, and σ_t used by the estimators (with respect to the true values in the measurements) by adding a normally-distributed error with zero mean and standard deviation equal to 10% of the true values — which means a maximum error of about 30%. Results showed that the proposed ad-hoc estimator is extremely robust to such mismatches, with deviations that are not appreciable in the graphic. We considered a second set of simulations where the perturbation has been increased to 20% of the true values (i.e., maximum error of about 60% considering three standard deviations) and a non-zero mean equal to 10% of σ_t has been added in the TOA. The latter is meant to emulate non-perfect calibration and other factors that may produce residual errors in this respect. Results reported in Fig. 9 clearly show that also in this case the proposed estimator yields superior performance despite some deviation from the case of perfect calibration, which however is not greater than the one of the ML benchmark.

As final remark, notice that by design the proposed relaxation approach "shrinks" the estimator in a certain interval depending on the values of the control points $\{d_1, \ldots, d_n\}$ in $\hat{d}_{\text{JAH}}$, or more simply on $[A, B]$ if a uniformly-spaced sampling is chosen (see Sec. IV). This allows one to tune the behavior according to the scenario at hand. Just to make an example, in the case considered above if additional information is available to expect d will lie within a more restricted interval $[\ell, u]$, a rule of thumb is to set $A = \ell - 5$ and $B = u + 25$. By using this heuristic, one can optimize the performance of $\hat{d}_{\text{JAH}}$ over the interval $[\ell, u]$. Results are reported in Fig. 10 for different intervals and values of σ: it is apparent that $\hat{d}_{\text{JAH}}$ can outperform $\hat{d}_{\text{JML}}$ in the chosen range, paying the price with a larger error elsewhere (shaded area).



(a) $[\ell, u] = [4, 12]$



(b) $[\ell, u] = [8, 18]$

Fig. 10. Performance improvement of $\hat{d}_{\text{JAH}}$ for a restricted span.

VI. Conclusions

We have addressed the problem of ranging based on hybrid TOA/RSS measurements. We have performed a thorough theoretical analysis, showing that in many cases of practical interest for WSN applications the inter-node separation significantly overlaps with the region around the critical distance, where hybrid TOA/RSS estimation is superior to TOA-only and RSS-only ranging. To circumvent the need for iterative optimization to compute the joint ML estimator, we proposed a suitable relaxation that ultimately leads to a closed-form ad-hoc estimator amenable to practical implementation on low-cost devices. The performance assessment revealed that the proposed solution outperforms the natural competitor [30], and may be even superior to ML estimation, especially if coarse information (bound) is available to restrict the distance span of interest. As a whole, the proposed approach is general and is shown to be robust also to mismatches in the calibration of the model parameters.

Appendix

A. Proof of Proposition 1

The main tools used in the proof are the implicit function theorem, the Taylor expansion, and the chain rule for differentiation. Our approach follows the line of [50], but on a rather different case. Eq. (9) defines a function $h(\boldsymbol{y})$ that maps the data $\boldsymbol{y}$ into $\hat{d}_{\text{JML}}$. In our case an analytical expression for $h(\boldsymbol{y})$ is unknown (the function is implicit); moreover, we are interested in computing some moments (to obtain bias

and variance/MSE) which is even harder due to the involved nonlinearity. We can however consider the Taylor expansion of $h(\boldsymbol{y})$, i.e.

$$h(\boldsymbol{y}) \approx h(\bar{\boldsymbol{y}}) + \frac{\partial h(\bar{\boldsymbol{y}})}{\partial r}(r-\bar{r}) + \frac{\partial h(\bar{\boldsymbol{y}})}{\partial t}(t-\bar{t}) + \frac{1}{2}\left[\frac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial r^2}(r-\bar{r})^2 + \frac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial r \partial t}(r-\bar{r})(t-\bar{t}) + \frac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial t \partial r}(r-\bar{r})(t-\bar{t}) + \frac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial t^2}(t-\bar{t})^2\right]$$

where the sign $\approx$ indicates that h.o.t. have been neglected, and $\bar{x}$ denotes the mean value of x. Thus

$$\mathrm{E}[\hat{d}_{\mathrm{JML}}] \approx h(\bar{\boldsymbol{y}}) + \frac{1}{2}\frac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial r^2}\sigma_r^2 + \frac{1}{2}\frac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial t^2}\sigma_t^2 \tag{27}$$

where we have exploited independence between r and t. Since $\bar{\boldsymbol{y}} = \mathrm{E}[\boldsymbol{y}] = \begin{bmatrix} P_0 - a\log d \\ T_0 + \frac{d}{c} \end{bmatrix}$, it turns out that $h(\bar{\boldsymbol{y}}) = \arg\min L(\bar{\boldsymbol{y}}) = d$, i.e., the ML procedure returns the true value d because $\bar{\boldsymbol{y}}$ corresponds to noise-free measurements. The computation of the mean still requires knowledge of the second derivatives of h; however, what is needed is only the value in $\boldsymbol{y} = \bar{\boldsymbol{y}}$. To this aim, we preliminary differentiate (9) with respect to (w.r.t.) r and t by applying the chain rule, obtaining

$$\begin{cases} L_{d^2}\dfrac{\partial h(\bar{\boldsymbol{y}})}{\partial r} + L_{dr} = 0 \\ L_{d^2}\dfrac{\partial h(\bar{\boldsymbol{y}})}{\partial t} + L_{dt} = 0 \end{cases} \tag{28}$$

where, introducing the notation $L_{d^p r^q t^s} = \left.\frac{\partial^3 L(h(\boldsymbol{y}),\boldsymbol{y})}{\partial d^p \partial r^q \partial t^s}\right|_{\boldsymbol{y}=\bar{\boldsymbol{y}}}$ (subscripts omitted for order 0), it turns out that $L_{d^2} = \frac{2(a^2\sigma_t^2 c^2 + \sigma_r^2 d^2)}{c^2 d^2}$, $L_{dr} = \frac{2\sigma_t^2 a}{d}$, $L_{dt} = -\frac{2\sigma_r^2}{c}$, with $a \stackrel{\text{def}}{=} \frac{10}{\ln 10}\alpha$. Then, chain-differentiating (28) w.r.t. r and t, we can solve the system for the two unknowns $\frac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial r^2}$ and $\frac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial t^2}$, as follows:

$$\begin{cases} \dfrac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial r^2} = -\frac{1}{L_{d^2}}\left[(L_{d^3}h_r + L_{d^2r})h_r + L_{d^2r}h_r + L_{dr^2}\right] \\ \dfrac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial t^2} = -\frac{1}{L_{d^2}}\left[(L_{d^3}h_t + L_{d^2t})h_t + L_{d^2t}h_t + L_{dt^2}\right] \end{cases}$$

where $L_{d^3} = -\frac{2a\sigma_t^2(2a+1)}{d^3}$, $L_{d^2r} = -\frac{2\sigma_t^2}{d^2}$, $L_{dr^2} = L_{d^2t} = L_{dt^2} = 0$, therefore

$$\begin{cases} \dfrac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial r^2} = \dfrac{L_{dr}}{L_{d^2}^2}(L_{d^3}h_r + 2L_{d^2r}) \\ \dfrac{\partial^2 h(\bar{\boldsymbol{y}})}{\partial t^2} = \dfrac{L_{dt}}{L_{d^2}^2}L_{d^3}h_t \end{cases} \tag{29}$$

since

$$h_r = \frac{\partial h(\bar{\boldsymbol{y}})}{\partial r} = -\frac{L_{dr}}{L_{d^2}}, \quad h_t = \frac{\partial h(\bar{\boldsymbol{y}})}{\partial t} = -\frac{L_{dt}}{L_{d^2}}. \tag{30}$$

Substituting into (29) yields the second-order derivatives to be used into (27), hence the thesis.

B. Proof of Proposition 2

Starting from the Taylor expansion as in Appendix A

$$h(\boldsymbol{y}) \approx h(\bar{\boldsymbol{y}}) + \frac{\partial h(\bar{\boldsymbol{y}})}{\partial r}(r-\bar{r}) + \frac{\partial h(\bar{\boldsymbol{y}})}{\partial t}(t-\bar{t})$$

and taking the VAR[$\cdot$] operator at both sides, we obtain

$$\mathrm{VAR}[\hat{d}_{\mathrm{JML}}] \approx h_r^2\sigma_r^2 + h_t^2\sigma_t^2$$

where independence between r and t and variance properties are exploited. Using (30), the thesis follows by recalling that $\mathrm{MSE}[\hat{d}_{\mathrm{JML}}] = \mathrm{VAR}[\hat{d}_{\mathrm{JML}}] + \left(\mathrm{E}[\hat{d}_{\mathrm{JML}}] - d\right)^2$.

C. Proof of Proposition 3

The well-known solution of problem (22), which can be obtained by taking derivatives, is given by

$$m^\star = \frac{\sum_{i=1}^n (d_i - \bar{d})(z_i - \bar{z})}{\sum_{i=1}^n (d_i - \bar{d})^2}, \qquad q^\star = \bar{z} - m^\star\bar{d}$$

where $\bar{d}$ and $\bar{z}$ denote the arithmetic means of d_is and z_is, respectively, i.e., $\bar{d} \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n d_i}{n}$ and $\bar{z} \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n z_i}{n}$. Using such expressions, eq. (23) can be rewritten as

$$\mathrm{e}^{m^\star d} = \mathrm{e}^{-q^\star} d. \tag{31}$$

The solution of the transcendental equation (31) can be expressed using the Lambert W-function. The latter is defined (in the complex plane) as the inverse function of $f(w) = w\mathrm{e}^w$; since $f(\cdot)$ is not injective, W is in general a multivalued function. If $m^\star$ is negative, we can consider the *principal branch* W_0 (starting at $w \geq -1$) [51] and write $\hat{d}_{\mathrm{JAH}} = -\frac{1}{m^\star}W_0\left(-m^\star\mathrm{e}^{q^\star}\right), m^\star < 0$, which is a well-defined estimator taking on values in $[0, +\infty)$ due to the fact that $W(x) \geq 0$ for $x \geq 0$ (monotonically increasing from $W(0) = 0$) [52].

Recalling that $z_i = z(d_i; \boldsymbol{y})$ and exploiting (21), the expressions $z_i = \eta_i t - \frac{r}{a} + \theta_i$ and $\bar{z} = t\bar{\eta} - \frac{r}{a} + \bar{\theta}$ are obtained, where $\eta_i = \frac{\gamma^2 c}{a^2}d_i$, $\theta_i = \frac{P_0}{a} - \frac{\gamma^2}{a^2}d_i^2 - \frac{\gamma^2 c T_0}{a^2}d_i$, and $\bar{\eta}$ and $\bar{\theta}$ are the corresponding arithmetic means. A tedious calculation yields the equivalent expression $m^\star = \eta t + \theta$, where $\eta = \frac{\gamma^2 c}{a^2}$ and θ is a function of $\delta_k = \frac{1}{n}\sum_{i=1}^n d_i^k$, $k = 1, 2, 3$. Further calculations lead to $m^\star = \frac{\gamma^2}{a^2}[c(t - T_0) - \delta]$ where $\delta = \frac{\delta_3 - \delta_1\delta_2}{\delta_2 - \delta_1^2}$; by proceeding analogously, we get $q^\star = \frac{P_0 - r}{a} - \frac{\gamma^2}{a^2}(\delta_2 - \delta_1\delta)$. Posing $\varphi \stackrel{\text{def}}{=} -m^\star$ and $\psi \stackrel{\text{def}}{=} q^\star$ the thesis follows straight.

The expression (24) needs to be slightly modified if $c(t - T_0) > \delta$, i.e., $m^\star > 0 \Leftrightarrow \varphi < 0$. In this case the Lambert W-function has two branches (W_0 and W_{-1}) due to its negative argument $\varphi\mathrm{e}^\psi$, until the point $-1/\mathrm{e}$, then the function turns complex-valued [48], [49], [52]. Since $\varphi\mathrm{e}^\psi \geq -1/\mathrm{e} \Leftrightarrow \varphi \geq -\frac{1}{\mathrm{e}^{\psi+1}}$, (23) has two solutions for $\varphi \in (-\frac{1}{\mathrm{e}^{\psi+1}}, 0)$ and no solution for smaller values. This is easily interpreted: while the curves $\log(d)$ and $z(d; \boldsymbol{y})$ will always intersect in $\hat{d}_{\mathrm{JML}}$, a positive-slope line $-\varphi d + \psi$ may intersect $\log d$ either in two points or nowhere. In the former case, one takes the solution with the minimum value of L; in the latter case, since $W(-1/\mathrm{e}) = -1$, by continuity argument the estimator can be clipped to $-\frac{1}{\varphi}$.

D. Proof of Proposition 4

Eq. (25) can be rewritten as $d_i = p_0 + p_1\, i$ with $p_0 = A - \frac{B-A}{n-1} = A - p_1$ and $p_1 = \frac{B-A}{n-1}$. From the formulas for the sum of the first n integers, squares, and cubes: $\delta_1 = p_0 + \frac{p_1}{2}(n+1)$, $\delta_2 = p_0^2 + p_0 p_1 (n+1) + \frac{p_1^2}{6}(n+1)(2n+1)$, $\delta_3 = p_0^3 + \frac{3}{2}p_0^2 p_1 (n+1) + \frac{p_0 p_1^2}{2}(n+1)(2n+1) + \frac{p_1^3}{n}\left(\frac{n(n+1)}{2}\right)^2$. Substituting in the numerator and denominator of δ, after a tedious calculation one obtains $\delta_3 - \delta_1\delta_2 = \frac{p_1^2}{12}(n+1)(n-1)\,[2p_0 + p_1(n+1)]^2$ and $\delta_2 - \delta_1^2 = \frac{p_1^2}{12}(n+1)\,[4n+2-3n-3]$, thus $\delta = 2A - 2p_1 + p_1 n + p_1 = A + B$ and $\tilde{\delta}$ follows.

E. Proof of Proposition 5

The proof is similar to the one of Proposition 1, but with an important difference: while an analytical expression for $h(\boldsymbol{y})$ is unknown (the function is implicit), $\hat{d}_{\text{JAH}}$ has a closed form. We thus consider the Taylor expansion of the equation in Proposition 3; similarly to (27) in Appendix A

$$\mathrm{E}[\hat{d}_{\text{JAH}}] \approx \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}}) + \frac{1}{2}\frac{\partial^2 \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}})}{\partial r^2}\sigma_r^2 + \frac{1}{2}\frac{\partial^2 \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}})}{\partial t^2}\sigma_t^2 \quad (32)$$

where again $\bar{\boldsymbol{y}} = \mathrm{E}[\boldsymbol{y}] = \begin{bmatrix} P_0 - a\log d \\ T_0 + \frac{d}{c} \end{bmatrix}$. Exploiting the fact that $\frac{\mathrm{d}W(x)}{\mathrm{d}x} = \frac{W(x)}{x(1+W(x))} = \frac{1}{x+\mathrm{e}^{W(x)}}$, a long calculation shows that the first and second partial derivatives of the function $\hat{d}_{\text{JAH}}$ in $\boldsymbol{y} = \bar{\boldsymbol{y}}$ are given by

$$\begin{aligned}
&\frac{\partial \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}})}{\partial r} = -\frac{1/a}{\lambda+\mu}, \quad \frac{\partial \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}})}{\partial t} = \frac{\gamma^2 c}{\mu a^2}\left(\xi - \frac{1}{\lambda+\mu}\right) \\
&\frac{\partial^2 \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}})}{\partial r^2} = -\frac{\lambda^2/a^2}{(\lambda+\mu)^3}, \\
&\frac{\partial^2 \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}})}{\partial t^2} = \left(\frac{\gamma^2 c}{a^2\mu}\right)^2\left[2\xi - \frac{1}{\lambda+\mu}\left(1 + \frac{2\mu\lambda+\mu^2}{(\lambda+\mu)^3}\right)\right] \quad (33)
\end{aligned}$$

where $\xi \stackrel{\text{def}}{=} \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}}) = \frac{1}{\mu} W_0\left(\mu\, d\, \mathrm{e}^{\frac{\gamma^2}{a^2}\tilde{\delta}}\right)$, $\mu \stackrel{\text{def}}{=} \varphi(\bar{\boldsymbol{y}}) = \frac{\gamma^2(\delta - d)}{a^2}$, and $\lambda \stackrel{\text{def}}{=} \mathrm{e}^{\varphi(\bar{\boldsymbol{y}})\hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}}) - \psi(\bar{\boldsymbol{y}})} = d\,\mathrm{e}^{\mu\xi - \frac{\gamma^2}{a^2}\tilde{\delta}}$. Since the bias is $\mathrm{E}[\hat{d}_{\text{JAH}}] - d$, substituting these expression in (32) finally yields the thesis.

F. Proof of Proposition 6

Starting from the Taylor expansion as in Appendix E,

$$\hat{d}_{\text{JAH}}(\boldsymbol{y}) \approx \left[\frac{\partial \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}})}{\partial r}\right]^2 \sigma_r^2 + \left[\frac{\partial \hat{d}_{\text{JAH}}(\bar{\boldsymbol{y}})}{\partial t}\right]^2 \sigma_t^2$$

and using the expressions in (33) the thesis follows by recalling that $\mathrm{MSE}[\hat{d}_{\text{JAH}}] = \mathrm{VAR}[\hat{d}_{\text{JAH}}] + \left(\mathrm{E}[\hat{d}_{\text{JAH}}] - d\right)^2$.

REFERENCES

[1] S. Gezici, "A survey on wireless position estimation," *Wireless Pers. Commun., Int. J.*, vol. 44, no. 3, pp. 263–282, Feb. 2008.

[2] M. Z. Win *et al.*, "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.

[3] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: Challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 24–40, Jul. 2005.

[4] W. Dai, Y. Shen, and M. Z. Win, "Energy-efficient network navigation algorithms," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1418–1430, Jul. 2015.

[5] A. Coluccia, F. Ricciato, and G. Ricci, "Positioning based on signals of opportunity," *IEEE Commun. Lett.*, vol. 18, no. 2, pp. 356–359, Feb. 2014.

[6] A. E. Waadt, C. Kocks, S. Wang, G. H. Bruck, and P. Jung, "Maximum likelihood localization estimation based on received signal strength," in *Proc. 3rd Int. Symp. Appl. Sci. Biomed. Commun. Technol. (ISABEL)*, Nov. 2010, pp. 1–5.

[7] A. Coluccia and F. Ricciato, "RSS-based localization via Bayesian ranging and iterative least squares positioning," *IEEE Commun. Lett.*, vol. 18, no. 5, pp. 873–876, May 2014.

[8] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.

[9] S. Tomic, M. Beko, R. Dinis, and P. Montezuma, "Distributed algorithm for target localization in wireless sensor networks using RSS and AoA measurements," *Pervasive Mobile Comput.*, vol. 37, pp. 63–77, Jun. 2017.

[10] S. Tomic, M. Beko, R. Dinis, and P. Montezuma, "A closed-form solution for RSS/AoA target localization by spherical coordinates conversion," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 680–683, Dec. 2016.

[11] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2137–2148, Aug. 2003.

[12] G. Mao, B. Fidan, and B. D. O. Anderson, "Wireless sensor network localization techniques," *Comput. Netw.*, vol. 51, no. 10, pp. 2529–2553, Jul. 2007.

[13] G. Wang and K. Yang, "A new approach to sensor node localization using RSS measurements in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1389–1395, May 2011.

[14] A. Coluccia and F. Ricciato, "On ML estimation for automatic RSS-based indoor localization," in *Proc. IEEE Int. Symp. Wireless Pervasive Comput.*, May 2010, pp. 495–502.

[15] S. Salari, S. Shahbazpanahi, and K. Ozdemir, "Mobility-aided wireless sensor network localization via semidefinite programming," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 5966–5978, Dec. 2013.

[16] S. Tiwari, D. Wang, M. Fattouche, and F. Ghannouchi, "A hybrid RSS/TOA method for 3D positioning in an indoor environment," *ISRN Signal Process.*, vol. 2012, 2012, Art. no. 503707. [Online]. Available: https://www.hindawi.com/journals/isrn/2012/503707/, doi: 10.5402/2012/503707.

[17] M. McGuire, K. N. Plataniotis, and A. N. Venetsanopoulos, "Data fusion of power and time measurements for mobile terminal location," *IEEE Trans. Mobile Comput.*, vol. 4, no. 2, pp. 142–153, Mar./Apr. 2005.

[18] J. Prieto, S. Mazuelas, A. Bahillo, P. Fernández, R. M. Lorenzo, and E. J. Abril, "Adaptive data fusion for wireless localization in harsh environments," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1585–1596, Apr. 2012.

[19] Y. Wang, F. Zheng, M. Wiemeler, W. Xiong, and T. Kaiser, "Reference selection for hybrid TOA/RSS linear least squares localization," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, Sep. 2013, pp. 1–5.

[20] Z. Sahinoglu and A. Catovic, "A hybrid location estimation scheme (H-LES) for partially synchronized wireless sensor networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2004, pp. 3797–3801.

[21] A. Bahillo, S. Mazuelas, J. Prieto, P. Fernández, R. M. Lorenzo, and E. J. Abril, "Hybrid RSS-RTT localization scheme for wireless networks," in *Proc. IEEE Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2010, pp. 1–7.

[22] M. Laaraiedh, S. Avrillon, and B. Uguen, "Hybrid data fusion techniques for localization in UWB networks," in *Proc. 6th IEEE Workshop Positioning, Navigat. Commun. (WPNC)*, Mar. 2009, pp. 51–57.

[23] M. Laaraiedh, L. Yu, S. Avrillon, and B. Uguen, "Comparison of hybrid localization schemes using RSSI, TOA, and TDOA," in *Proc. Eur. Wireless Conf.*, 2011, pp. 1–5.

[24] M. F. Keskin and S. Gezici, "Comparative theoretical analysis of distance estimation in visible light positioning systems," *J. Lightw. Technol.*, vol. 34, no. 3, pp. 854–865, Feb. 1, 2016.

[25] A. Catovic and Z. Sahinoglu, "The Cramér–Rao bounds of hybrid TOA/RSS and TDOA/RSS location estimation schemes," *IEEE Commun. Lett.*, vol. 8, no. 10, pp. 626–628, Oct. 2004.
[26] D. Macii, A. Colombo, P. Pivato, and D. Fontanelli, "A data fusion technique for wireless ranging performance improvement," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 1, pp. 27–37, Jan. 2013.
[27] Y. Qi and H. Kobayashi, "On relation among time delay and signal strength based geolocation methods," in *Proc. IEEE GLOBECOM*, Dec. 2003, pp. 4079–4083.
[28] Y. Wang, G. Leus, and A.-J. van der Veen, "Cramér–Rao bound for range estimation," in *Proc. IEEE ICASSP*, Apr. 2009, pp. 3301–3304.
[29] S. Sand, W. Wang, and A. Dammann, "Cramér–Rao lower bounds for hybrid distance estimation schemes," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, Sep. 2012, pp. 1–5.
[30] J. Zhang, L. Ding, Y. Wang, and L. Hu, "Measurement-based indoor NLoS ToA/RSS range error modelling," *Electron. Lett.*, vol. 52, no. 2, pp. 165–167, Jan. 2016.
[31] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.
[32] G. Zanca, F. Zorzi, A. Zanella, and M. Zorzi, "Experimental comparison of RSSI-based localization algorithms for indoor wireless sensor networks," in *Proc. ACM Workshop Real-World Wireless Sensor Netw.*, 2008, pp. 1–5.
[33] A. Coluccia, "Reduced-bias ML-based estimators with low complexity for self-calibrating RSS ranging," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1220–1230, Mar. 2013.
[34] H. Lim, L.-C. Kung, J. C. Hou, and H. Luo, "Zero-configuration, robust indoor localization: Theory and experimentation," in *Proc. IEEE INFOCOM*, Apr. 2006, pp. 1–12.
[35] A. Coluccia and F. Ricciato, "A software-defined radio tool for experimenting with RSS measurements in IEEE 802.15.4: Implementation and applications," *Int. J. Sensor Netw.*, vol. 14, no. 3, pp. 144–154. 2013.
[36] R. Bernhardt, "Macroscopic diversity in frequency reuse radio systems," *IEEE J. Sel. Areas Commun.*, vol. SAC-5, no. 5, pp. 862–870, Jun. 1987.
[37] F. Babich and G. Lombardi, "Statistical analysis and characterization of the indoor propagation channel," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 455–464, Mar. 2000.
[38] J. B. Andersen, T. S. Rappaport, and S. Yoshida, "Propagation measurements and models for wireless communications channels," *IEEE Commun. Mag.*, vol. 33, no. 1, pp. 42–49, Jan. 1995.
[39] U. Mengali and A.N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York, NY, USA: Springer, 1997.
[40] N. Patwari, R. J. O'Dea, and Y. Wang, "Relative location in wireless networks," in *Proc. IEEE Veh. Technol. Conf.*, May 2001, pp. 1149–1153.
[41] A. J. Weiss, "On the accuracy of a cellular location system based on RSS measurements," *IEEE Trans. Veh. Technol.*, vol. 52, no. 6, pp. 1508–1518, Nov. 2003.
[42] A. J. Weiss and J. S. Picard, "Network localization with biased range measurements," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 298–304, Jan. 2008.
[43] K. Doğançay, "On the bias of linear least squares algorithms for passive target localization," *Signal Process.*, vol. 84, no. 3, pp. 475–486, Mar. 2004.
[44] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Statistics), 2nd ed. New York, NY, USA: Springer, 2016.
[45] L. Breiman, "Statistical modeling: The two cultures," *Statist. Sci.*, vol. 16, no. 3, pp. 199–231, 2001.
[46] A. W. Bowman and A. Azzalini, *Applied Smoothing Techniques for Data Analysis*. London, U.K.: Oxford Univ. Press, 1997.
[47] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[48] D. A. Barry, P. J. Culligan-Hensley, and S. J. Barry, "Real values of the W-function," *ACM Trans. Math. Softw.*, vol. 21, no. 2, pp. 161–171, Jun. 1995.
[49] F. N. Fritsch, R. E. Shafer, and W. P. Crowley, "Solution of the transcendental equation $we^w = x$," *Commun. ACM*, vol. 16, no. 2, pp. 123–124, Feb. 1973.
[50] J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography," *IEEE Trans. Image Process.*, vol. 5, no. 3, pp. 493–506, Mar. 1996.
[51] N. G. de Bruijn, *Asymptotic Methods in Analysis*. New York, NY, USA: Dover, 1981.
[52] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert *W* function," *Adv. Comput. Math.*, vol. 5, no. 1, pp. 329–359, 1996.

Angelo Coluccia (M'13–SM'16) received the Eng. degree (*summa cum laude*) in telecommunication engineering and the Ph.D. degree in information engineering from the University of Salento, Lecce, Italy, in 2011 and 2007, respectively. Since 2008, he has been engaged in research projects on traffic analysis in operational cellular networks, security, and anomaly detection. He has been a Research Fellow with Forschungszentrum Telekommunikation Wien, Vienna, Austria, and a Visiting Scholar with the Institut Supérieur de l'A´eronautique et de l'Espace ISAE-Supaero, Toulouse, France. He is currently an Assistant Professor with the Dipartimento di Ingegneria dell'Innovazione, University of Salento, where he teaches the course of telecommunication systems. His research interests are in the area of multi-channel and multi-agent statistical signal processing, including cooperative sensing and estimation in wireless networks, detection, and localization. Relevant application fields are radar, ad-hoc (sensor, overlay, social) networks, and smart cities (including intelligent transportation systems).

Alessio Fascista was born in Nardó (Lecce), Italy, in 1991. He received the Eng. degree (*summa cum laude*) in computer engineering from the University of Salento, Lecce, Italy, in 2015. He is currently pursuing the Ph.D. degree in engineering of complex systems with the Department of Innovation Engineering, University of Salento. His main research interests are in the field of networking and telecommunications with focus on localization systems.