ROBUST BOOTSTRAP BASED OBSERVATION CLASSIFICATION FOR KALMAN FILTERING IN HARSH LOS/NLOS ENVIRONMENTS

Stefan Vlaski

Department of Electrical Engineering
University of California, Los Angeles
Box 951594, Los Angeles, CA 90095-1594, U.S.A.
Email: svlaski@ucla.edu

Abdelhak M. Zoubir

Signal Processing Group
Technische Universität Darmstadt
Merckstraße 25, 64283 Darmstadt, Germany
Email: zoubir@spg.tu-darmstadt.de

ABSTRACT

The bootstrap allows for the estimation of the distribution of an estimate without requiring assumptions on the distribution of the underlying data, relying on asymptotic results or theoretical derivations. In contrast to a point estimate, the distribution estimate captures the uncertainty about the statistic of interest. We introduce a novel robust bootstrap method and demonstrate how this additional information is utilized to improve the performance of robust tracking methods. A robust bootstrap method is crucial, because the classical bootstrap is highly sensitive to outliers, irrespective of the robustness of the underlying estimator. Using the robust distribution estimate of the state prediction as a measure of confidence, the bootstrap allows to incorporate an observation weighting scheme into the tracking algorithm, which enhances performance.

Index Terms— robust, bootstrap, confidence region, tracking, Extended Kalman Filter

1. INTRODUCTION

Tracking describes the problem of estimating the evolving state vector of a mobile terminal (MT) in a network of fixed terminals (FT). The Extended Kalman Filter (EKF) is a well established method for this task under a given, non-linear state space model. In practical methods, non-line-of-sight (NLOS) paths between the MT and FT's may occur which result in outliers from the LOS measurement model [1]. To mitigate these, robust versions of the EKF have been developed, that trade off efficiency in the LOS case and robustness in the NLOS case [2].

We propose an improvement to existing robust tracking approaches based on the bootstrap. For this, we consider Robust Extended Kalman Filtering, where at each time instance, a predicted estimate based on past observations is corrected using the newest set of observations. This correction can be large for two reasons:

1. The predicted estimate was uncertain. In this case, the estimate is greatly improved by the correction. This is desired.
2. The observation is corrupted due to NLOS outliers. In this case, the estimate is deteriorated by the correction. This is undesired.

If one could distinguish between the two cases and only correct the prediction if appropriate, the performance of the KF will be improved. Knowledge of the distribution of the predicted estimate would allow such inference because:

1. If the previous state estimate was uncertain, the predicted state estimate is uncertain. This results in increased variance of the distribution and large confidence regions. A correction within these regions is hence likely to be appropriate.
2. If the previous state estimate was certain, the predicted state will have a low variance distribution with tight confidence bounds. A correction outside of these regions is hence likely caused by NLOS outliers and inappropriate.

In practical applications, knowledge of these distributions is not given. The bootstrap is a tool that allows for their estimation [3]. The lack of robustness of the classical bootstrap was recognized and some robust bootstrap methods have been proposed ([4] and [5]). Recently, we introduced the Robust Starting Point Bootstrap (RSPB) and demonstrated its competitive performance to other methods [6]. It is based on the fact, that robustness properties and breakdown point of many robust estimators are determined by the initial starting point. Hence, by utilizing a robust starting point, robust estimates can be obtained even past the breakdown point of the estimator. The RSPB is applicable to a wide range of robust estimators, including M-, MM- and τ-estimators of regression.

The remainder of this paper is structured as follows: Sec. 2 introduces the Robust Starting Point Bootstrap (RSPB) in a general linear regression context. Sec. 3 formulates the tracking problem and model and briefly reviews Robust Extended Kalman Filtering. Sec. 4 describes how additional information from the bootstrap is used to improve the REKF. Sec. 5 presents simulations results and Sec. 6 concludes the paper.

2. ROBUST STARTING POINT BOOTSTRAP (RSPB)

The bootstrap is a powerful tool for statistical inference. For a given data sample χ of size N with fraction of contamination ε, a total of B bootstrap samples $\chi_i^*, i = 1, ..., B$ are obtained by sampling N times with replacement from χ. In expectation, ε_i^*, the level of contamination in χ_i^*, equals the level of contamination in the original sample. For small N and with increasing ε or B, the probability of obtaining bootstrap resamples with levels of contamination past the breakdown point of the underlying estimator increases. As a result, the classical bootstrap is highly sensitive to outliers, irrespective of the robustness properties of the estimator that is bootstrapped. Consider the general linear regression model

$$\boldsymbol{z} = \boldsymbol{D}\boldsymbol{s} + \boldsymbol{v}, \tag{1}$$

where $\boldsymbol{\chi} = (\boldsymbol{z}, \boldsymbol{D})$, $\boldsymbol{z} \in \mathbb{R}^{N\times 1}$, $\boldsymbol{D} \in \mathbb{R}^{N\times K}$ and $\boldsymbol{s} \in \mathbb{R}^{K\times 1}$ is the statistic of interest. $\boldsymbol{v} \in \mathbb{R}^{N\times 1}$ is the observation noise and is assumed to follow the ε contamination model, i.e. $(1-\varepsilon)\mathcal{N}(0, I_N\sigma_\varepsilon^2) + \varepsilon\mathcal{H}$, where $\mathcal{H}$ is an unspecified distribution. Robust estimation of $\boldsymbol{s}$ is achieved through robust regression estimators, such as the M-, τ- or MM-estimator. However, robustness for the original sample χ does not guarantee robustness for every bootstrap sample χ_i^* [6].

Robustly estimating $\boldsymbol{s}$ requires finding the absolute minimum of a non-convex cost function, which is determined by the estimator. The Robust Starting Point Bootstrap (RSPB) is based on the following observation: For every bootstrap sample χ^*, it is not desirable to find the absolute minimum of the cost function, but rather a local minimum close to $\hat{s}$, the estimate on χ.

Non-convex minimization problems can generally be divided in two steps:

- **Step 1.** Obtain a good candidate point or a number of candidate points.
- **Step 2.** Apply an iterative algorithm, which reduces the respective cost function until the estimate converges.

For the RSPB, **Step 1.** is replaced by **Step 1'.** for every bootstrap sample as follows:

- **Step 0.** Obtain $\hat{s}$ and the corresponding scale estimate $\hat{\sigma}$ for the original sample χ through the full algorithm described above. This is done only once prior to bootstrapping.
- **Step 1'.** For every bootstrap resample $\chi_i^*, i = 1, ..., B$, choose the estimate $\hat{s}$ with scale estimate $\hat{\sigma}$ from χ as the only candidate point for the minimization problem.
- **Step 2.** Apply an iterative algorithm, which reduces the respective loss function until the estimate converges to a solution of Eq. (1) for χ_i^* to obtain $\hat{s}_i^*$ and $\hat{\sigma}_i^*$.

By applying the RSPB, robust bootstrap replicates $\hat{s}_i^*$ are obtained, even if the bootstrap sample χ_i^* is contaminated past the breakdown point of the underlying estimator.

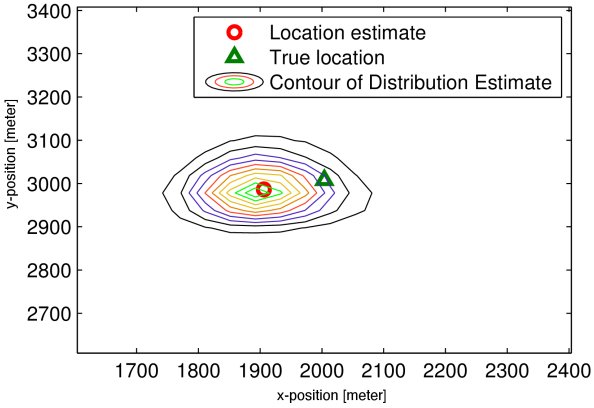


Fig. 1. RSPB distribution of $\hat{s}$, $\varepsilon = 0.4$, contour plot in 10% steps

Fig. 1 displays the contour plot of a distribution estimate obtained using the RSPB. Further details on the RSPB can be found in [6].

3. ROBUST TRACKING

We consider tracking of the state of a mobile terminal (MT) in a network of M fixed terminates (FT) with known locations $(x_{FT,i}, y_{FT,i}), i = 1, ..., M$. This is formulated in the following section. Throughout this paper, random variables are emphasized as bold letters, while their realizations are plain.

3.1. Problem Formulation

In this paper, we assume the state $\boldsymbol{s_n} = (\boldsymbol{x_n}, \boldsymbol{y_n}, \dot{\boldsymbol{x}}_{\boldsymbol{n}}, \dot{\boldsymbol{y}}_{\boldsymbol{n}})$ corresponding to the MT to be modeled by the state-space model with state equation

$$\boldsymbol{s_n} = A\boldsymbol{s_{n-1}} + G\boldsymbol{u_{n-1}}, \tag{2}$$

where A and G are transition and mapping matrices respectively according to

$$A = \begin{bmatrix} I_2 & T_s \cdot I_2 \\ 0 & I_2 \end{bmatrix}, \qquad G = \begin{bmatrix} T_s/2 \cdot I_2 \\ T_s \cdot I_2 \end{bmatrix}.$$

T_s is the sampling period and I_D denotes the $D \times D$ identity matrix. The driving noise $\boldsymbol{u_n}$ is assumed to be white with zero mean. The state is then related to the observations by the measurement equation:

$$\boldsymbol{z_n} = h(\boldsymbol{s_n}) + \boldsymbol{v_n} \tag{3}$$

Here, $h(\boldsymbol{s_n})$ is a real-valued, non-linear, vector-valued function. For time of arrival (TOA) estimation,

$$h_i(\boldsymbol{s_n}) = \sqrt{(\boldsymbol{x_n} - x_{FT,i})^2 + (\boldsymbol{y_n} - y_{FT,i})^2}, i = 1, ..., M$$

is the Euclidean distance between the position of the MT and each FT [1]. $\boldsymbol{v_n}$ is the measurement noise with covariance matrix R_n.

3.2. Extended Kalman Filter

An approach to deal with the non-linear nature of $h(\boldsymbol{s_n})$ is the Extended Kalman Filter (EKF). It involves a first-order Taylor approximation of $h(\boldsymbol{s_n})$ and results in [1]:

$$\begin{aligned} e_n &= z_n - h(\hat{s}_{n|n-1}) \\ R_{e,n} &= H_n P_{n|n-1} H_n^T + R_n \\ K_n &= P_{n|n-1} H_n^T R_{e,n}^{-1} \\ \hat{s}_{n|n} &= \hat{s}_{n|n-1} + K_n e_n \\ P_{n|n} &= (I_4 - K_n H_n) P_{n|n-1} \end{aligned}$$

Here, e_n are the innovations and $\hat{s}_{n|n-1}$ is the state estimate at time n from observations up to time $n-1$. $R_{e,n}$ is the innovation covariance matrix, K_n is the Kalman gain and $P_{n|n}$ is the error covariance matrix. H_n is the Jacobian matrix of partial derivatives, evaluated at the predicted state estimate $\hat{s}_{n|n-1} = A\hat{s}_{n-1|n-1}$.

3.3. Robust Extended Kalman Filter

The EKF performs well as long as the observation noise $\boldsymbol{v_n}$ follows a distribution with known covariance matrix R_n. In practical applications, high-mean, high-variance errors are caused by non-line-of-sight (NLOS) paths between the MT and FT's and result in deviations from the model. Hence, a robust EKF (REKF) is required. A common approach to robustifying the EKF is to rewrite the state and observation equations (Eq. (2) and (3)) into a single linear regression model

$$\tilde{\boldsymbol{z}}_{\boldsymbol{n}} = \boldsymbol{D}\boldsymbol{s_n} + \tilde{\boldsymbol{v}}_{\boldsymbol{n}}$$

and apply robust linear regression techniques [7]. In this paper, the semi-parametric M-estimator, introduced in [1], is used. By passing the estimate from the original set of observation as starting point to the M-estimator for each bootstrap resample, the RSPB is directly applicable.

4. RSPB AND THE REKF

The RSPB allows for the robust estimation of the distribution of the state estimate $\hat{s}_{n|n}$ and the corresponding state prediction distribution of $\hat{s}_{n+1|n}$ by applying the following procedure. At every time instant n,

1. Observe z_n and find the corresponding innovations e_n.
2. Obtain B bootstrap samples $e^*_{n,i}, i = 1, ..., B$ by sampling from the elements of the innovation vector e_n M times with replacement.
3. Run the Robust Extended Kalman procedure on each bootstrap sample to obtain B state estimates $\hat{s}^*_{n|n,i}, i = 1, ..., B$.
4. Obtain the empirical state prediction distribution according to $\hat{s}^*_{n+1|n,i} = A\hat{s}^*_{n|n,i}, i = 1, ..., B$.
5. Estimate the distribution of $\hat{s}_{n+1|n}$ by partitioning the x-y-plane into I intervals of equal size and counting the number of $\hat{s}^*_{n+1|n,i}$ in each interval. This results in an estimate of the probability density function (pdf).

Fig. 2 illustrates the motivation for combining the RSPB and the REKF. The previous and predicted state are depicted along with their respective bootstrap distributions. It is observed, that the corrected estimate highly deviates from the predicted estimate as well as its distribution. It can be inferred, that the correction results from NLOS outliers and is hence inappropriate.

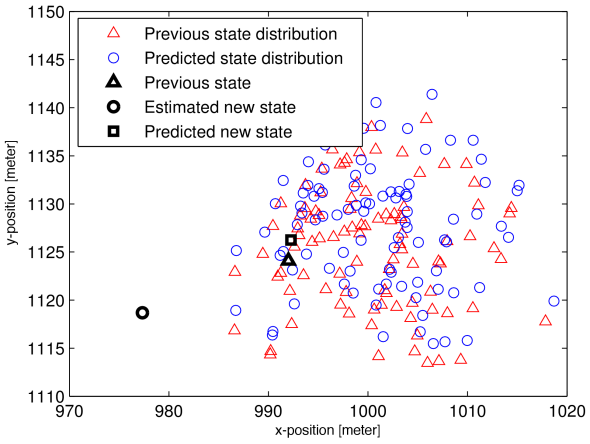


Fig. 2. Concept Illustration

The full algorithm is then summarized as follows. At each time instance n,

1. Observe z_n.
2. **REKF**: Obtain the REKF estimate $\hat{s}_{n|n,\mathrm{REKF}}$ by correcting the state prediction $\hat{s}_{n|n-1} = A\hat{s}_{n-1|n-1}$ using the innovations e_n.
3. **Bootstrap**: Obtain bootstrap replicates $\hat{s}^*_{n|n,\mathrm{REKF}}, i = 1, ..., B$ by bootstrapping from the elements of e_n.
4. **Confidence Evaluation** Evaluate the confidence c_n based on the state prediction distribution estimate $f_{n|n-1}$ from the previous iteration.

5. **State Estimate Correction**: Obtain the linear combination

$$\hat{s}_{n|n} = c_n \cdot \hat{s}_{n|n,\mathrm{REKF}} + (1 - c_n) \cdot \hat{s}_{n|n-1}.$$

6. **Distribution Estimate Correction**: Obtain the corrected bootstrap distribution estimate by

$$\hat{s}^*_{n|n} = c_n \cdot \hat{s}^*_{n|n,\mathrm{REKF}} + (1 - c_n) \cdot \hat{s}_{n|n-1}, i = 1, ..., B$$

and the state prediction bootstrap replicates $\hat{s}^*_{n+1|n} = A\hat{s}^*_{n|n}, i = 1, ..., B$. Estimate $f_{n+1|n}$ which is used in the next iteration.

By comparing the corrected state estimate to the state prediction distribution in every iteration, NLOS heavy observations are identified and weighed down accordingly. This results in a more robust tracking procedure.

5. SIMULATION RESULTS

Simulations are carried comparing the performance of the EKF, REKF and Bootstrapped REKF in a mixed LOS-NLOS environment. The robust M-estimator for both the REKF and the Bootstrapped REKF is set with parameters $c_1 = 1.5, c_2 = 3.5$. A MT with initial state $s_0 = (1000, 1000, 3, 3)$ moves in a network of $M = 5$ FT's located at $(500, 1500)$, $(2500, 1500)$, $(1500, 2500)$, $(1500, 500)$ and $(1500, 1500)$. The sampling interval is $T_s = 0.2$ and $N = 1000$ time instances are observed over 100 MC iterations. Fig. 3 displays the empirical cumulative density function of the Euclidean distances between the position estimate and the true position of the MT. The occurrence of the outliers is modeled by a Markov chain and the value of the outliers follows an exponential distribution with $\sigma_{\mathrm{NLOS}} = 300$ and $\varepsilon = 0.5$. For the purpose of quantifying the relationship between the new state estimate and the predicted state distribution, $B = 100$ and $I = 5$ were found to be appropriate.

The simulation results show, that the information about the uncertainty of a state prediction captured by the bootstrap can indeed be utilized to improve the performance of the REKF in harsh LOS/NLOS environments.

6. CONCLUSION

The Robust Starting Point Bootstrap, which we developed for the linear regression model was applied to the tracking of a MT. It was demonstrated, how additional information about the uncertainty of an estimate is utilized to softly classify incoming observations and improve the performance of the Robust Extended Kalman Filter. High corrections to a state prediction can be caused by uncertain predictions or heavy NLOS contamination. The bootstrap can be utilized to softly classify and weigh accordingly. The application to more sophisticated robust tracking methods, such as interactive multiple model algorithms or parametric approaches is subject to future research.

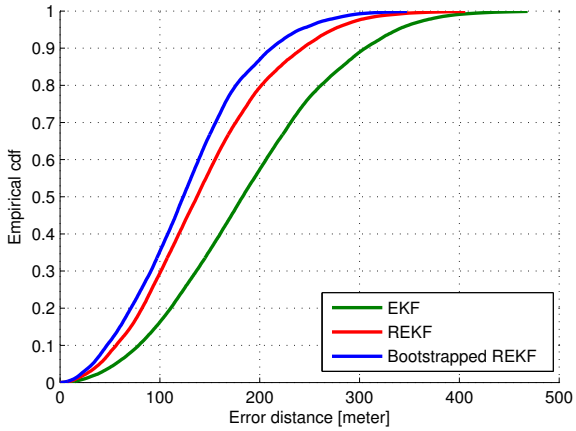


Fig. 3. Empirical cumulative density function of positioning errors, $\varepsilon = 0.5$

7. REFERENCES

[1] U. Hammes and A.M. Zoubir, "Robust MT Tracking Based on M-Estimation and Interacting Multiple Model Algorithm," *Signal Processing, IEEE Transactions on*, vol. 59, no. 7, pp. 3398–3409, July 2011.

[2] T. Perala and R. Piche, "Robust Extended Kalman Filtering in Hybrid Positioning Applications," in *Positioning, Navigation and Communication, 2007. WPNC '07. 4th Workshop on*, March 2007, pp. 55–63.

[3] A. M. Zoubir and D. R. Iskander, *Bootstrap Techniques for Signal Processing*, Cambridge University Press, 2004.

[4] S. Müller and A. H. Welsh, "Outlier Robust Model Selection in Linear Regression," *Journal of the American Statistical Association*, vol. 100, no. 472, pp. 1297–1310, 2005.

[5] M. Salibian-Barrera and R. H. Zamar, "Bootstrapping Robust Estimates of Regression," *The Annals of Statistics*, vol. 30, no. 2, pp. 556–582, 2002.

[6] S. Vlaski, M. Muma, and A. M. Zoubir, "Robust Bootstrap Methods with an Application to Geolocation in Harsh LOS/NLOS Environments," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP'14)*, May 2014, accepted.

[7] P. Huber and E. M. Ronchetti, *Robust Statistics*, John Wiley & Sons, 2nd edition, 2009.