The Annals of Statistics
2015, Vol. 43, No. 4, 1448–1454
DOI: 10.1214/15-AOS1270D
Main article DOI: 10.1214/14-AOS1270

DISCUSSION OF "FREQUENTIST COVERAGE OF ADAPTIVE NONPARAMETRIC BAYESIAN CREDIBLE SETS"[1]

BY MARK G. LOW AND ZONGMING MA

University of Pennsylvania

We congratulate the authors for this very interesting article focused on the frequentist coverage of Bayesian credible sets in the context of an infinite-dimensional signal in white noise models. In such settings the construction of honest confidence sets is especially complicated, at least when the goal is to construct confidence sets that have a size that adapts to the unknown parameters in the model, while maintaining coverage probability.

The focus of the present paper is on constructing l_2 balls as confidence sets. There are some advantages that come with the focus on balls for confidence sets. For bands results in Low (1997) rule out the possibility of adaptation over even a pair of Lipschitz or Sobolev spaces at least for confidence bands that have a guaranteed coverage level. On the other hand, fully rate adaptive confidence balls which do maintain coverage probability can be constructed over Sobolev smoothness levels that range over an interval $[\alpha, 2\alpha]$. However, this range of models where such adaptation is possible is still quite limited and here the authors develop a theory that applies over a broader class of models. The approach taken, following Giné and Nickl (2010) and Bull (2012), is to focus on parameters that are in some sense typical and removing a set of parameter values that cause difficulties at least when constructing adaptive sets. The goal is then to construct fully adaptive confidence sets over the remaining collection of parameter values. In the present paper the parameter values that are kept belong to a class of parameters that they call polished tail sequences and the authors develop results that show that a particular empirical Bayes credible ball is both honest when restricted to such sequences and adaptive in size.

There are of course many settings where it is more natural to focus on the construction of confidence bands rather than confidence balls and, typically, theory and methodology developed for balls do not provide a way to also construct bands. Here, however, the balls are constructed from an empirical Bayes posterior and even though the focus of the paper is on the construction of balls, the simulation example in Section 4 suggests that a general methodology for the construction of confidence bands can also be developed based on this posterior. The visualization of the credible sets is constructed by making draws from the empirical Bayes posterior and plotting the 95% that are closest in l_2 to the posterior mean. Each draw gives rise to an entire function, but visually the appearance is somewhat akin to a

Received January 2015.
[1]Supported in part by NSF Grants DMS-13-52060 and DMS-14-03708.

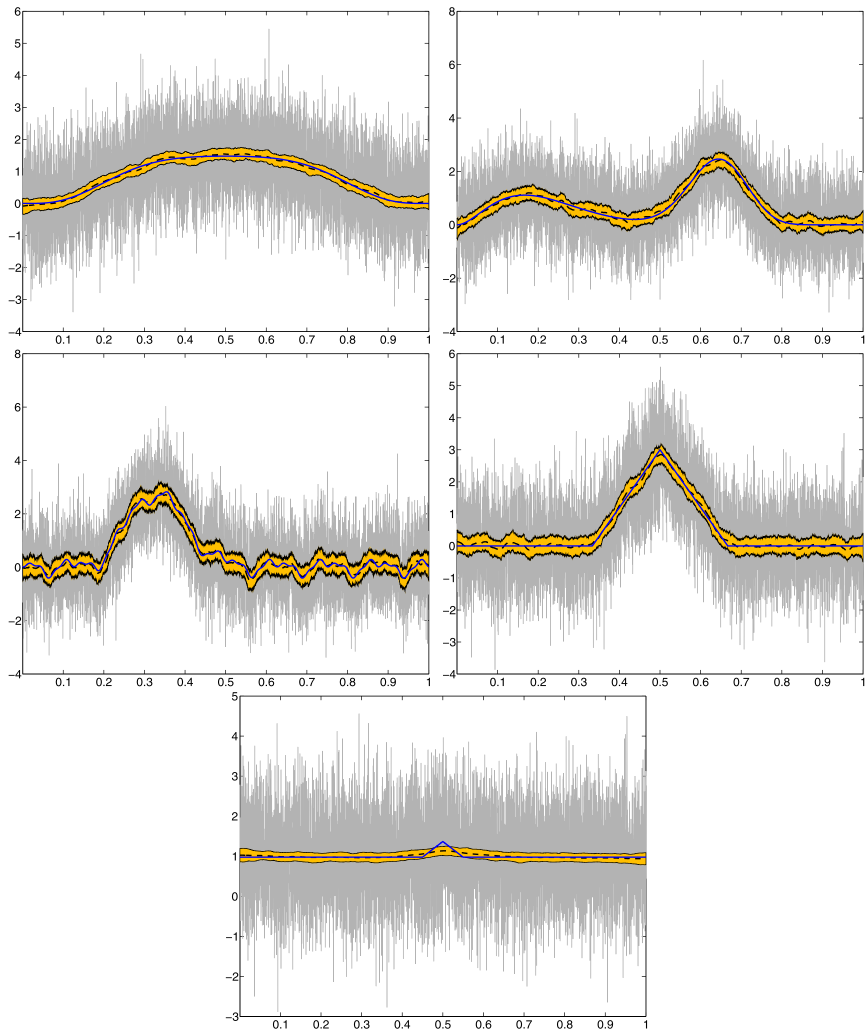

FIG. 1. *One realization of the observed data with* $n = 8192$ *observed points and the resulting band for the EBayes procedure*. *Case* 1: *top left*; *case* 2: *top right*; *case* 3: *middle left*; *case* 4: *middle right*; *case* 5: *bottom*. *Black solid*: *the true function*. *Gray*: *observed data*. *Orange*: *confidence band*. *Black dashed*: *band center*.

confidence band and claims from the picture of good coverage could perhaps also be interpreted from that point of view.

Looking closely at the pictures in Figure 1, it appears that either the entire function is covered or there is only a very small region where the true function is not covered by such a credible set. Although the visualizations given in Figure 1 are not technically bands, it is quite easy to make true bands as follows. First generate N realizations from the posterior and keep the 95% that are closest in l_2 to the mean. This gives a collection of curves, $f_1, f_2, \dots, f_m$ where $m = 0.95N$. A band

$[L(t), U(t)]$ can then be made by taking pointwise the max and min of these functions, $L(t) = \min_{1\le i\le m} f_i(t)$ and $U(t) = \max_{1\le i\le m} f_i(t)$.

In this discussion, we explore this approach in the context of the nonparametric regression model

$$y_i = f(t_i) + \sigma\varepsilon_i, \qquad i = 1, \dots, n, \tag{1}$$

where $t_i = \frac{i}{n}$ and $\varepsilon_i \overset{\text{i.i.d.}}{\sim} N(0,1)$, and make some comparisons with bands found in Cai, Low and Ma (2014). As mentioned above, truly adaptive bands do not exist over the most commonly considered function spaces. Cai, Low and Ma (2014) develop a new formulation for such problems by relaxation of the requirement that the entire function is covered by the confidence band. Two approaches are considered. In the first the goal is to minimize the expected width of the confidence band while maintaining coverage at most of the points in [0, 1] where the expected width adjusts to the smoothness of the underlying function. The second approach is to limit the excess mass of the function lying outside the confidence band while once again minimizing the expected width of the confidence bands.

We report here how the proposed confidence band based on the empirical Bayes posterior performs in terms of this new formulation and compare the performance with the adaptive confidence band procedure considered in Cai, Low and Ma (2014). We consider five test functions. The first four of these were also considered in Cai, Low and Ma (2014) and three of these were considered earlier in Wahba (1983). The five functions are as follows:

Case 1. $f(t) \propto B_{10,5}(t) + B_{7,7}(t) + B_{5,10}(t)$,
Case 2. $f(t) \propto 3B_{30,17}(t) + 2B_{3,11}(t)$,
Case 3. $f(t) \propto 7B_{15,30}(t) + 2\sin(32\pi t - \frac{2\pi}{3}) - 3\cos(16\pi t) - \cos(64\pi t)$,
Case 4 $f(t) \propto (t - \frac{1}{3})I(\frac{1}{3} \le t \le \frac{1}{2}) + (\frac{2}{3} - t)I(\frac{1}{2} \le t \le \frac{2}{3})$,
Case 5 $f(t) \propto 1 + 8(t - 0.45)I(0.45 \le t \le 0.5) + 8(0.55 - t)I(0.5 \le t \le 0.55)$,

where $B_{a,b}(t)$ stands for the density function of a Beta(a,b) distribution. In all cases, we rescale the function so that $\int_0^1 f^2 = 1$ and we take $\sigma = 1$.

In order to construct the band based on the above empirical Bayes posterior approach, we first apply a discrete cosine transform to the regression data. This yields the observations $X_j = \frac{1}{n}\sum_{i=1}^n y_i \cos((j - \frac{1}{2})\pi t_i)$. The observation sequence $X = (X_1, X_2, \dots)$ then satisfies equation (2.1) of the present paper with $\kappa_i = 1$. It is then easy to construct the confidence bands based on the empirical Bayes posterior as suggested above. Note that the construction is not entirely automatic, as the number of draws N from the posterior needs to be specified. The number of draws for the empirical Bayes (EBayes) band cannot be taken too large, or the band will be very wide, or too small because then the band has little hope of covering the unknown function. However, in the examples given below we found that for values of N that ranged from 2000 to 20,000, the width of the interval grew by only

TABLE 1
Simulation results from 500 *repetitions*: f_1, *each with* 2000 *posterior draws*

	ACB				
#Sample	**Mean size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**	
1024	1.326	0.002	<0.001	0.924	
8192	0.446	0	0	0.990	

	EBayes					
#Sample	**Mean max size**	**Mean ave size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**	L_2 **coverage**
1024	1.246	0.858	0	0	0.990	1.000
8192	0.520	0.358	0	0	0.986	0.996

around 15% and, thus, from a purely methodological point of view, the method does not appear too sensitive to the choice of this parameter. In the simulation results given below we take $N = 2000$, the same value that is used in the paper to generate the pictures from the simulations from the empirical Bayes procedure.

For the adaptive confidence band (ACB) there are two parameters that need to be chosen. The choice of these parameters results in control of the set of noncovered points as well as control of the excess mass over a collection of smoothness classes. In the experiments given below we always take $\beta_0 = 2$ and $M_0 = 1000$, and in this case the adaptation results that are given in Cai, Low and Ma (2014) are for a range of smoothness between 2 and 4. Of course, in practice, it is not always clear whether a function would belong to a particular smoothness class and both case 4 and case 5 fall outside the range.

In Tables 1–5 we report the mean width of the adaptive confidence band procedure found in Cai, Low and Ma (2014). Figure 2 shows representative realizations

TABLE 2
Simulation results from 500 *repetitions*: f_2, *each with* 2000 *posterior draws*

	ACB				
#Sample	**Mean size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**	
1024	1.969	0.003	<0.001	0.918	
8192	0.673	0	0	0.980	

	EBayes					
#Sample	**Mean max size**	**Mean ave size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**	L_2 **coverage**
1024	1.783	1.299	0	0	0.990	1.000
8192	0.792	0.536	0	0	0.978	1.000

TABLE 3
Simulation results from 500 *repetitions*: f_3, *each with* 2000 *posterior draws*

	ACB			
#Sample	**Mean size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**
1024	1.965	0.005	<0.001	0.888
8192	0.911	0.003	<0.001	0.932

	EBayes					
#Sample	**Mean max size**	**Mean ave size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**	L_2 **coverage**
1024	2.083	1.442	0	0	0.974	1.000
8192	1.048	0.707	<0.001	<0.001	0.934	1.000

TABLE 4
Simulation results from 500 *repetitions*: f_4, *each with* 2000 *posterior draws*

	ACB			
#Sample	**Mean size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**
1024	1.337	0.003	<0.001	0.912
8192	0.669	0	0	0.990

	EBayes					
#Sample	**Mean max size**	**Mean ave size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**	L_2 **coverage**
1024	1.859	1.278	0	0	0.978	1.000
8192	0.826	0.558	0	0	0.952	1.000

TABLE 5
Simulation results from 500 *repetitions*: f_5, *each with* 2000 *posterior draws*

	ACB			
#Sample	**Mean size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**
1024	1.341	0.003	< 0.001	0.912
8192	0.459	0.024	0.003	0.298

	EBayes					
#Sample	**Mean max size**	**Mean ave size**	$NC_{0.95}$	$RE_{0.95}$	L_∞ **coverage**	L_2 **coverage**
1024	0.841	0.588	0.033	0.004	0.552	0.980
8192	0.419	0.388	0.039	0.009	0.318	0.878

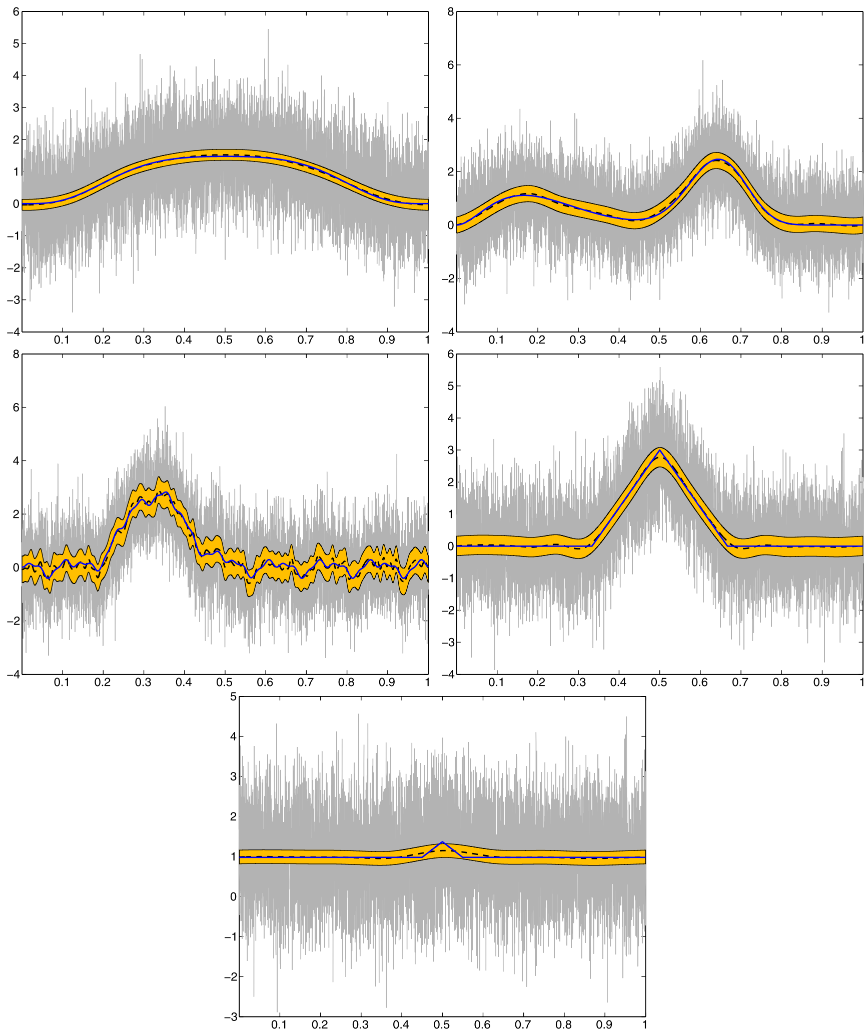

FIG. 2. *One realization of the observed data with* $n = 8192$ *observed points and the resulting band for the ACB procedure. Case* 1: *top left*; *case* 2: *top right*; *case* 3: *middle left*; *case* 4: *middle right*; *case* 5: *bottom. Black solid*: *the true function. Gray*: *observed data. Orange*: *confidence band. Black dashed*: *band center.*

of the band on the five test functions. Although the width of the band is random for a given set of data, it has fixed width over the interval. The EBayes band is of variable width and we report both the mean maximum width and then the mean average width. For each replication we also calculated the fraction of the interval where the function is not covered as well as the relative excess mass, and we report the 95th percentiles of these values based on 500 replications. Finally, we also report the fraction of the time that the bands cover the whole function and also, in the case of the EBayes procedure, the coverage of the associated L_2 balls.

For each of these test functions we find that the EBayes procedure performs quite well from the point of view of the framework given in Cai, Low and Ma (2014).

REFERENCES

BULL, A. D. (2012). Honest adaptive confidence bands and self-similar functions. *Electron*. *J*. *Stat*. **6** 1490–1516. MR2988456

CAI, T. T., LOW, M. and MA, Z. (2014). Adaptive confidence bands for nonparametric regression functions. *J*. *Amer*. *Statist*. *Assoc*. **109** 1054–1070. MR3265680

GINÉ, E. and NICKL, R. (2010). Confidence bands in density estimation. *Ann*. *Statist*. **38** 1122–1170. MR2604707

LOW, M. G. (1997). On nonparametric confidence intervals. *Ann*. *Statist*. **25** 2547–2554. MR1604412

WAHBA, G. (1983). Bayesian "confidence intervals" for the cross-validated smoothing spline. *J*. *Roy*. *Statist*. *Soc*. *Ser*. *B* **45** 133–150. MR0701084

DEPARTMENT OF STATISTICS
THE WHARTON SCHOOL
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PENNSYLVANIA 19104
USA
E-MAIL: zongming.ma@gmail.com