When calculations/summations are analyzed it is because Naive Bayes only includes the alpha probability of the given label, whereas Naive Bayes EM includes alpha probability of each label. Thus, Naive Bayes EM also includes the log probability of either the predicted probability or the given probability for each label. Interpreting calculations, the fact of uncertainty of some labels require the likelihood function to consider probabilities for possible outcomes rather than only taking account of the given outcomes. Which results in a lower likelihood.

The reason why accuracy is not positively correlated to likelihood is likely because for the parameters(alpha and beta) when the model produces 100% accuracy it's probably also producing high confidence probabilities for unknown labels. Which is not realistic and we don't want the model to provide near 100% or near 0% probabilities for unknown labels. Hence our likelihood calculation penalizes that and the model iteratively adjusts probabilities to not provide very high confidence results for unknown labels. Hence, the adjusted parameters lead to misclassification for some documents leading to a decrease in accuracy, but to a better/more realistic overall performance/distribution.