BAYESIAN INVERSION OF MULTI-MODE NEMS MASS SPECTROMETRY SIGNAL

R. Pérenon[1], E. Sage[1], A. Mohammad-Djafari[2], L. Duraffourg[1],
S. Hentz[1], A. Brenac[3], R. Morel[3], P. Grangeat[1].

[1]CEA Leti, Minatec Campus
17 Rue des Martyrs, F38054 Grenoble Cedex 9, France
[2]Laboratoire des Signaux et Systèmes (L2S), UMR 8506 (CNRS-SUPELEC-UNIV. PARIS SUD)
Supélec, Plateau de Moulon, 3 Rue Joliot Curie, 91192 Gif-sur-Yvette, France
[3]INAC/SP2M and Université Joseph Fourier, CEA Grenoble
17 Rue des Martyrs, F38054 Grenoble Cedex 9, France

ABSTRACT

Nano ElectroMechanical Systems are a new class of sensors that offers high sensitivity and opens new perspectives in the mass spectrometry field. This acquisition is performed in counting-mode, and the main tasks of associated information processing are to detect the molecules, to quantify their respective mass and to combine this information in order to recover the mass spectrum of the analysed solution.

We propose a joint detection-quantification method based on a hierarchical description of the measurement system. Computation is done using a Reversible Jumps Monte-Carlo Markov-Chain algorithm.

The approach we are describing in this communication solves the two problems of the joint impulse deconvolution on multiple output signals (multi-mode acquisition) and the non-linear relation between the observed signals and the mass of molecules, including the localization of the molecules on the sensor. We validate our method on both simulated and experimental data.

Index Terms— Inverse problems, Information processing, Statistical signal processing, Mass spectrometry, Bayesian inference, NEMS, Nanotechnologies, MCMC, Proteomics, Detection-Estimation

1. INTRODUCTION

Mass spectrometry is a research field aiming at estimating the molecular profile of a given solution. Many technologies enables this estimation, such as Time-of-Flight (ToF), ion traps or quadrupole selection [1]. One particularly significant application of mass spectrometry is the detection of cancers in the early stage [2]. Related information processing issues are typically the estimation of a molecular profile based on device output signal, the search of discriminant molecules for a particular disease, or the classification of samples [3]. Conventional Mass Spectrometry instruments are all based on a flow-mode principle, which means that the molecules are ionized to get an ion flow which is then separated and detected.

New sensors, called Nano ElectroMechanical Systems (NEMS), are currently developed in CEA-Leti in collaboration with the California Institute of Technology within the NanoVLSI alliance [4]. These sensors are sensitive enough to detect one single molecule in the mass range of a protein (hundreds of kDa) by tracking shifts of resonant frequency of the sensor over time. The first proof of concept of NEMS mass spectrometry was given by Naik *et al.* [5]. We previously described a Bayesian method to estimate resonant frequency shifts in the mono-mode acquisition case [6]. This communication extends this previous method to a multi-mode acquisition, enabling the estimation of a mass spectrum.

2. PROBLEM STATEMENT

2.1. Measurement principle

A typical device used for Mass Spectrometry is given in Figure 1. To track resonant frequency over time, the NEMS resonator (yellow part in Figure 1) is inserted into an electronic system containing an electrostatic actuator (blue part in Figure 1), in order to excite the sensor and to create an in-plane motion, and two piezoresistive gauges (red part in Figure 1), in order to detect the position of the sensor.

The physical principle of molecules detection has been summarized by Dohn *et al.* [7]. The NEMS sensor can be described as a resonator with a resonant frequency f_0 given by $f_0 = \sqrt{\frac{s}{M_0}}$, where s denotes the resonator spring constant and M_0 denotes the effective mass of the sensor. The resonator presents several vibration modes. Therefore, it has several resonant frequencies, named harmonics, and written
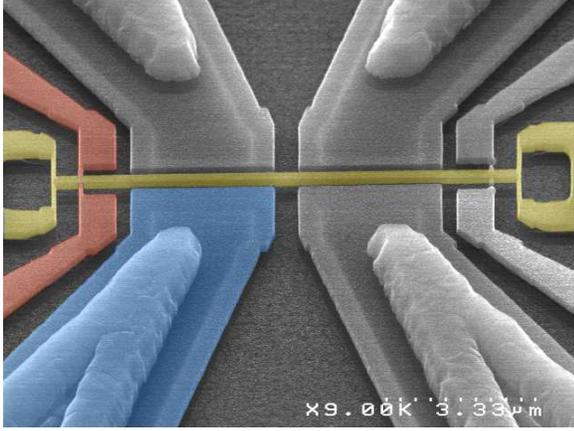
Fig. 1. Colourized scanning electron microscope image of a NEMS sensor

f_k for the k-th harmonic.

Let us suppose that a molecule of mass δm lands on the sensor at the normalized position z (varying between 0 and 1). To first order approximation and under the assumption that the resonator spring constant s is not affected by the mass addition, the new resonant frequency $f_{k,\delta m}$ on a specific harmonic k, can be written as:

$$f_{k,\delta m} = f_{k,0} - \alpha_k \, \delta m \, \phi_k^2(z) \tag{1}$$

In this equation, α_k is the constant gain on the k-th harmonic and the function ϕ_k^2 represents the variable gain on the k-th harmonic depending on the molecule position. Expression of the non-linear functions ϕ_k is:

$$\begin{aligned} \phi_k(z) = & \; A_k \left(\cosh(\kappa_k z) - \cos(\kappa_k z) \right) \\ + & \; B_k \left(\sinh(\kappa_k z) - \sin(\kappa_k z) \right) \end{aligned} \tag{2}$$

These functions are symmetric for a double-clamped geometry sensor, as the one used in this paper. Thus, the position of the adsorbed molecule can indifferently be z or $1 - z$.

2.2. Observed signals model

In the previous subsection, we showed that the adsorption of a molecule on the sensor leads to a drop of resonant frequency on numerous harmonics, and that the amplitude of the frequency shift depends on the mass and the position of the adsorbed molecule. To monitor the sensor resonant frequency over time, the NEMS sensor is included into an electronic system based on a actuator, a detector of sensor position and a phase-lock loop (PLL) electronic circuit (one PLL circuit per harmonic observed). Under the assumption that the response time of the PLL is significantly longer than the physical resonator's response time to an adsorption, the elementary signal $g_{e\,k}(t)$ induced on the k-th harmonic by the adsorption of a

molecule of mass δm at the position z is:

$$g_{e\,k}(t) = g_k(0) - \alpha_k \, \delta m \, \phi_k^2(z) \, h_k(t) \tag{3}$$

where $h_k(t)$ is the response of the electronic system to a normalized sharp change in resonant frequency of the sensor.

Now, let us consider the sequential adsorption of numerous molecules. The mass of molecule with index i is written m_i, its position z_i and its adsorption time t_i. The output signal $g_k(t)$ is the addition of every elementary signals:

$$g_k(t) = g_k(0) - \sum_{i=1}^{N} \alpha_k \, m_i \, \phi_k^2(z_i) \, h_k(t - t_i) \tag{4}$$

The constant gain values in our case are $\alpha_1 = 0.0498$ and $\alpha_2 = 0.1255$ if the mass and the resonant frequency are respectively expressed in kDa and in Hz.

The overall output signal $g_k(t)$ is also affected by noise, whose origins are multiple [8] and will not be discussed here. The noise contribution is modelled in section 3.3.

The signal is digitalized on T samples. These digital signals (one per harmonic), constituting the system output, are smooth and contains fast decreasing changes synchronized on every signals, corresponding to the adsorption of a molecule on the sensor. In our case, the experiments were performed using two harmonics.

The information processing aims at detecting frequency changes and estimating the mass of the landed molecules. Then, this information can be combined to estimate the mass spectrum of the initial solution. Multi-mode signal acquisition and first constitution of a NEMS Mass Spectrometry spectrum are notably described by Hanay *et al.* [9].

3. BAYESIAN APPROACH

3.1. Motivations

We introduced the forward model in the previous section, *i.e.* the physical equations linking the unknown quantities to the observable ones (the signals). We propose here to use an inverse problem approach associated with the Bayesian framework to recover the unknown quantities. We assume that the parameters α_k and the functions $h_k(t)$ and $\phi_k(z)$ are known. The unknown parameters are the number of adsorbed molecules N, the adsorption times $\{t_i\}$, the mass of the adsorbed molecules $\{m_i\}$ and their respective positions $\{z_i\}$.

The expression of the forward problem is not sufficient here. Indeed, without additional information, the inverse problem is ill-posed in the sense of Hadamard [10]. The problem must be regularized. Tikhonov [11] or Lasso [12] are classical regularization schemes based on penalization.

We propose to use the Bayesian framework which offers numerous advantages. Notably, this framework allows an automatic computation of the hyperparameters, provides more robust solutions and gives ways to understand penalization

methods. Besides, a link exists between Maximum A Posteriori (MAP) Bayesian estimation and penalization [13].

3.2. Bayesian estimation and prior distributions

The proposed Bayesian estimation framework is based on the following probabilistic modelling of the observed signal $g_k(t)$ and the unknown parameters of the forward model (4). Refer to [14] for further information about the statistical distributions used in this communication.

The observed signals $g_k(t)$ depends on the following unknown parameters N, $\{t_i\}$, $\{m_i\}$, $\{z_i\}$. The noise is parametrized by covariance matrix denoted as $\boldsymbol{\Sigma}_k$ and is discussed in section 3.3.

Our Bayesian model is derivated from the Bernoulli-Gaussian model [15]. The number of adsorbed molecules, N, is considered as the sum of independent Bernoulli trials. Thus, we assign a Binomial prior on N. The parameter of success probability (or counting intensity) is denoted as π.

The π parameter also will be estimated by the Bayesian algorithm. We assign the conjugate prior distribution, consisting in a Beta distribution and denoted as $\mathcal{B}(\pi|a_\pi, b_\pi)$. (a_π, b_π) are π prior distribution hyperparameters and will be set to express prior information on intensity rate.

For each adsorption time $\{t_i\}$ we assign an uniform prior on the acquisition period, denoted as $\mathcal{U}_{[0,\,T]}(t_i)$ to express our lack of knowledge on adsorption time. In the same way for each adsorption normalized location $\{z_i\}$ we assign an uniform prior between 0 (the extremity of the beam) and 0.5 (the middle of the beam), denoted as $\mathcal{U}_{[0,\,0.5]}(z_i)$.

In the case of the mass of adsorbed molecules $\{m_i\}$, we assign a Gamma prior, to ensure their positivity, denoted as $\mathcal{G}(m_i|k_m, \theta_m)$. (k_m, θ_m) are $\{m_i\}$ prior distribution hyperparameters and will be set to express prior information on the mass distribution of the analysed mixture.

This model can be summed up with the graphical model described in Figure 2.
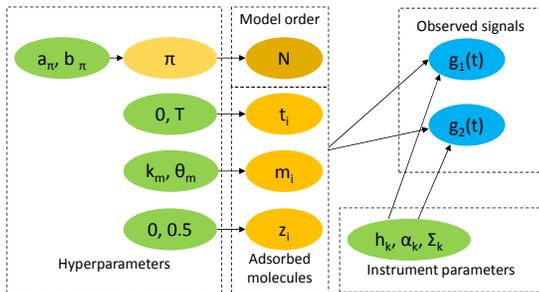
Fig. 2. Hierarchical model of the observed signals

3.3. Noise component

Incomplete knowledge of the noise components is translated into the following approximations:

- We consider that the noise is additive, and independent on each harmonic signals.

- We model the noise on each harmonic signal with a centred multivariate normal probability density function. On each k-th harmonic signal, the noise covariance matrix is denoted as $\boldsymbol{\Sigma}_k$.

The crucial point is the structure of the covariance matrix. Considering the noise is merely flicker noise (1/f noise), we assume that the covariance matrix can be written as $\boldsymbol{\Sigma}_k = \sigma_k^2 \mathbf{D}^{-1}\mathbf{D}^{-T}$ where σ_k represents the noise level on the k-th harmonic signal, and $\mathbf{D}$ is the derivation matrix. The values for σ_k in our case are: $\sigma_1 = 0.87$ and $\sigma_2 = 3.29$.

3.4. *A posteriori* distribution

From the hierarchical model in the previous section, we can deduce the posterior distribution of all unknown parameters. We denote as $\mathcal{D}(\cdot)$ the forward model described in equation (4). For a given model dimension N, the posterior of $\{t_i\}$, $\{m_i\}$, $\{z_i\}$ and π parameters is:

$$
\begin{aligned}
&p\left(t_1, \ldots t_N, m_1, \ldots m_N, z_1, \ldots z_N, \pi | N, g_1, g_2\right) \\
\propto\ &\sum_{k=1}^{2} \mathcal{N}\left(g_k | \mathcal{D}\left(t_1, \ldots t_N, m_1, \ldots m_N, z_1, \ldots z_N\right), \boldsymbol{\Sigma}_k\right) \\
&\left[\prod_{i=1}^{N} \mathcal{U}_{[0,\,T]}(t_i)\right]\left[\prod_{i=1}^{N} \mathcal{G}\left(m_i | k_m, \theta_m\right)\right] \\
&\left[\prod_{i=1}^{N} \mathcal{U}_{[0,\,0.5]}(z_i)\right] \mathcal{B}\left(\pi | a_\pi + N, b_\pi + T - N\right)
\end{aligned}
\tag{5}
$$

4. INVERSION ALGORITHM

4.1. Structure

We propose in this paper to combine the Reversible-Jumps Monte-Carlo Markov Chain (RJMCMC) and the Monte-Carlo Markov Chain (MCMC) algorithms [14]. In this paragraph, we describe the overall structure of our algorithm, which is constituted of two loops, described in next paragraphs. The first loop consists in a RJMCMC algorithm and chooses the model dimension N according to a Marginalized Maximum A Posteriori (MMAP) choice and $\{t_i\}$ according to a Maximum A Posteriori choice. The second loop consists in a MCMC algorithm and estimates the $\{m_i\}$ and $\{z_i\}$ parameters based on Expectation A Posteriori (EAP) choice.

4.2. First loop - RJMCMC algorithm

The first loop consists in a model choice algorithm. Each iteration is constituted of two parts. The first part is the sampling of the model with the RJMCMC algorithm. The evolution of the model order is very close from the Single-Most Likely Replacement algorithm described by Idier [15]. It consists in proposing the addition or the suppression of an adsorption. If a new adsorption is proposed, the new adsorption time, the

mass of the molecule and the position of adsorption are sampled according to their *a priori* distributions. If a suppression is proposed, the adsorption number (the i index) is uniformly chosen. This procedure is expressed as a derivation of the Green algorithm [14] to compute the acceptance probability.

The second part is the sampling of the parameters within the model according to the MCMC principle. π is sampled with its conditional posterior probability which is a Beta distribution. Each $\{t_i\}$ are sampled with a uniform random walk (one step forward or backward move), each $\{m_i\}$ and $\{z_i\}$ are sampled with a normal random walk (with standard deviations respectively δ_m and δ_z).

This algorithm is computed for N_{i1} iterations. After this, the most often sampled model $\widehat{N}$ from iteration $B_1 + 1$ to N_{i1} (the first B_1 values are discarded) is chosen. The $\widehat{N}$ most often sampled adsorption times are chosen as $\{\widehat{t_i}\}$ estimation.

4.3. Second loop - MCMC algorithm

The second loop consists in a MCMC algorithm, corresponding to the second part of the RJMCMC algorithm. Nevertheless, we do not need to sample π since $\widehat{N}$ is chosen. We also do not sample $\{t_i\}$ to prevent label switching. Indeed, if two $\{t_i\}$ are switching, the corresponding masses and positions can also switch. Each $\{m_i\}$ and each $\{z_i\}$ are sampled as described previously with a normal random walk.

This algorithm is computed for N_{i2} iterations. Then, the masses and the positions are estimated with:

$$\widehat{m_i} = \sum_{l=B_2+1}^{N_{i2}} \frac{m_i^{(l)}}{N_{i2} - B_2} \quad \widehat{z_i} = \sum_{l=B_2+1}^{N_{i2}} \frac{z_i^{(l)}}{N_{i2} - B_2} \quad (6)$$

In these equations, B_2 is the number of initial discarded values, $m_i^{(l)}$ and $z_i^{(l)}$ are respectively the Markov chains of the i-th estimated mass and the i-th estimated position.

5. RESULTS

5.1. Results on experimental data

Our algorithm is tested on experimental data acquired at CEA Grenoble. These data were acquired with the previously described NEMS sensor, inserted into a low-pressure chamber where tantalum nano-clusters were projected onto the NEMS surface. A time-of-flight mass spectrometer gave us the diameter spectrum of these nano-aggregates. The mass spectrum was centred on 6.5 nm diameter. We consider an observed signal g of 5000 samples, for a 500 Hz sampling frequency.

With the prior parameters and the algorithm parameters given in Table 1, the algorithm estimates the frequency shifts given in Figure 3 (represented by vertical arrows). All visible frequency shifts were estimated. We also compute the diameter d of every detected nano-clusters, given in Table 2. The estimated diameters mean is 6.53 nm, which is consistent with the mass spectrum of the analysed nano-clusters.

k_m	θ_m	a_π	b_π	N_{i1} N_{i2}	B_1 B_2	δ_m	δ_z
2	1000	1	5000	10000	8000	50	0.1

Table 1. Parameters used for experimental data processing
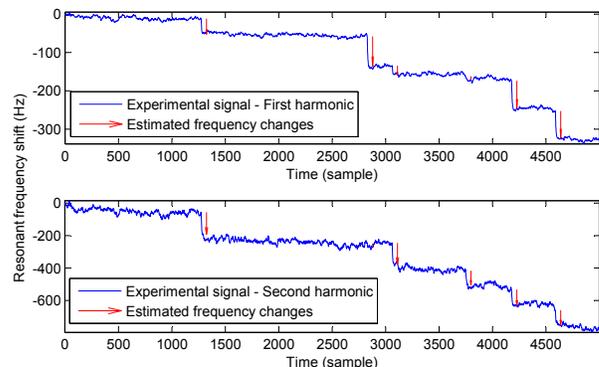
Fig. 3. Estimation results on an experimental signal

Time	1274	2828	3061	3748	4178	4588
d (nm)	6.26	6.78	5.88	6.58	6.90	6.80

Table 2. Table of estimated diameters on experimental data

5.2. Results on simulated data

To quantify more precisely the performances of our algorithm in term of detection capability and mass quantification, and especially to process noisier data, we propose to work on simulated data. The data are simulated using the model described in the section 2.2. We generate signals with 10 adsorbed molecules on 1000 samples, which is a higher counting-rate than the experimental one. The mass of molecules are sampled with a normal prior centred on M_0 with a standard deviation of $\frac{M_0}{20}$. The positions and the adsorption times are uniformly sampled. We generate data for numerous signal levels. The noise model and the standard deviation values σ_1 and σ_2 used are the same as in section 3.3. We define the Maximal Signal to Noise Ratio (MSNR) as $MSNR = M_0 \frac{\alpha_1 + \alpha_2}{\sigma_1 + \sigma_2}$.

The prior parameters and the algorithm parameters are set to values given in Table 3. The method is compared with two algorithms based on the deconvolution of the two signals separately with the pseudo-inverse method. The first algorithm, named averaged pseudo-inverse, consists in a deconvolution step, an averaging of these two deconvoluted signals and a thresholding step. The second one, named conservative pseudo-inverse, consists in a deconvolution step, a thresholding of both deconvoluted signal and an "AND" operation.

We compute the specificity of the algorithm (one minus the ratio of wrong detections over non-events number) and the sensitivity (ratio of true detections over events number). The thresholds in pseudo-inverse methods are designed to get a 99.9% specificity since our algorithm presents a minimal 99.9% specificity for all studied MSNR. The sensitivity re-

sults are given in Figure 4. We can see that our method outperforms pseudo-inverse ones.

k_m	θ_m	a_π	b_π	N_{i1} N_{i2}	B_1 B_2	δ_m	δ_z
2	200	1	100	10000	8000	50	0.1

Table 3. Parameters used for simulated data processing
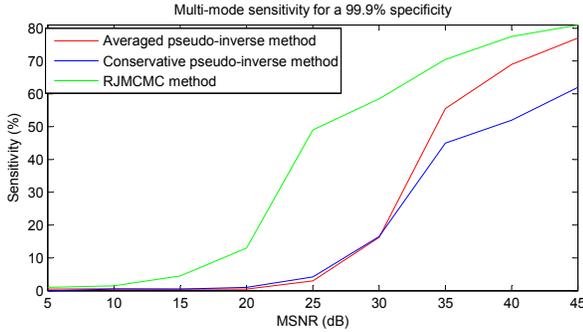
Fig. 4. Evolution of sensitivity on simulated data

We also compute the evolution of the Coefficient of Variation of the Root Mean Square Deviation (CV_{RMSD}) on mass estimation with the MSNR, given in Table 4. The CV_{RMSD} decreases with the mass of molecules and remains below 20% until the nominal mass of 421 kDa.

MSNR (dB)	35	30	25	20	15	10
Mass (kDa)	1332	749	421	236	133	75
CV$_{RMSD}$(%)	15.5	16.7	18.8	41.5	85.2	99.5

Table 4. Results of mass quantification on simulated data

6. CONCLUSION

In this article, we sum up the principle of NEMS Mass Spectrometry. We express the forward model since we model the acquisition system. The two specific points here are impulse deconvolution on several signals simultaneously and the inversion of a non-linear function. We compute the inversion with the Bayesian framework associated with an hybrid RJMCMC-MCMC algorithm. We test our method on experimental and simulated data. The results show that our algorithm outperforms a classical deconvolution method.

7. REFERENCES

[1] R. Aebersold and M. Mann, "Mass spectrometry-based proteomics," *Nature*, vol. 422, no. 6928, pp. 198–207, 2003.

[2] J. D. Wulfkuhle, L. A. Liotta, and E. F. Petricoin, "Proteomic applications for the early detection of cancer," *Nature Reviews Cancer*, vol. 3, no. 4, pp. 267–275, 2003.

[3] P. Szacherski, J. . Giovannelli, and P. Grangeat, "Joint bayesian hierarchical inversion-classification and application in proteomics," in *IEEE Workshop on Statistical Signal Processing Proceedings*, 2011, pp. 121–124.

[4] E. Mile, G. Jourdan, I. Bargatin, S. Labarthe, C. Marcoux, P. Andreucci, S. Hentz, C. Kharrat, E. Colinet, and L. Duraffourg, "In-plane nanoelectromechanical resonators based on silicon nanowire piezoresistive detection," *Nanotechnology*, vol. 21, no. 16, 2010.

[5] A. K. Naik, M. S. Hanay, W. K. Hiebert, X. L. Feng, and M. L. Roukes, "Towards single-molecule nanomechanical mass spectrometry," *Nature Nanotechnology*, vol. 4, no. 7, pp. 445–450, 2009.

[6] R. Pérenon, A. Mohammad-Djafari, E. Sage, L. Duraffourg, S. Hentz, A. Brenac, R. Morel, and P. Grangeat, "MCMC-based inversion algorithm dedicated to NEMS mass spectrometry," *Paper to appear in the* AIP *Conference Proceedings*, 2012.

[7] S. Dohn, W. Svendsen, A. Boisen, and O. Hansen, "Mass and position determination of attached particles on cantilever based mass sensors," *Review of Scientific Instruments*, vol. 78, no. 10, 2007.

[8] A. N. Cleland and M. L. Roukes, "Noise processes in nanomechanical resonators," *Journal of Applied Physics*, vol. 92, no. 5, pp. 2758–2769, 2002.

[9] M. S. Hanay, S. Kelber, A. K. Naik, D. Chi, S. Hentz, E. C. Bullard, E. Colinet, L. Duraffourg, and M. L. Roukes, "Single-protein nanomechanical mass spectrometry in real time," *Nature Nanotechnology*, vol. 7, no. 9, pp. 602–608, 2012.

[10] J. Hadamard, "Sur les problmes aux drives partielles et leur signication physique," *Princeton University Bulletin*, vol. 13, pp. 4952, 1902.

[11] A. Tikhonov, "Regularization of incorrectly posed problems," *Soviet. Math. Dokl.*, vol. 4, pp. 16241627, 1963.

[12] R. Tibshirani, "Regression shrinkage and selection via the lasso: A retrospective," *Journal of the Royal Statistical Society.Series B: Statistical Methodology*, vol. 73, no. 3, pp. 273–282, 2011.

[13] A. Mohammad-Djafari, "Bayesian approach with prior models which enforce sparsity in signal and image processing," *Eurasip Journal on Advances in Signal Processing*, vol. 2012, no. 1, 2012.

[14] C. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer Texts in Statistics. Springer, 2004.

[15] J. Idier, *Approche bayésienne pour les problmes inverses*, Herms Science Publications, Paris, 2001, ISBN 2-7462-0348-0.