Pre-whitening and Null Projection as an Artifact Suppression Method for Electrocorticography Stimulation in Bi-Directional Brain Computer Interfaces

Jeffrey Lim[1], Po T. Wang[1], Susan J. Shaw[2], Michelle Armacost[2], Hui Gong[2], Charles Y. Liu[2], An H. Do[3], Payam Heydari[4] and Zoran Nenadic[1,4]

***Abstract*— Electrocorticography (ECoG)-based bi-directional (BD) brain-computer interfaces (BCIs) are a forthcoming technology promising to help restore function to those with motor and sensory deficits. A major problem with this paradigm is that the cortical stimulation necessary to elicit artificial sensation creates strong electrical artifacts that can disrupt BCI operation by saturating recording amplifiers or obscuring useful neural signal. Even with state-of-the-art hardware artifact suppression methods, robust signal processing techniques are still required to suppress residual artifacts that are present at the digital back-end. Herein we demonstrate the effectiveness of a pre-whitening and null projection artifact suppression method using ECoG data recorded during a clinical neurostimulation procedure. Our method achieved a maximum artifact suppression of 21.49 dB and significantly increased the number of artifact-free frequencies in the frequency domain. This performance surpasses that of a more traditional independent component analysis methodology, while retaining a reduced complexity and increased computational efficiency.**

I. INTRODUCTION

Electrocorticography (ECoG)-based brain-computer interfaces (BCIs) have shown promise in the treatment of individuals with motor deficits such as quadriplegia [1] or amyotrophic lateral sclerosis (ALS) [2]. Additionally, subdural ECoG electrodes can elicit somatosensation by delivering cortical electrostimulation [3], [4]. Thus, it may be possible to add somatosensory capabilities to existing ECoG-based BCIs, which primarily rely on visual feedback [1], [2]. Interfaces with a closed somatosensory loop should be capable of restoring both motor and sensory functions and are often referred to as bi-directional (BD) BCIs [5]. Preliminary studies with intracortical microelectrodes show that somatosensory feedback may also improve BCI performance [6].

Enabling BD-BCI operation, however, requires several technical problems to be solved. For example, cortical electrostimulation creates strong electrical artifacts that propagate from the stimulation site to the recording site. These artifacts can undermine the function of the BD-BCI system as they can saturate recording front-ends as well as obfuscate neural information.

Supported by National Science Foundation (Award # 1446908, 1646275)

[1]Department of Biomedical Engineering, University of California Irvine (UCI), Irvine, CA 92697, USA {limj4,znenadic}@uci.edu

[2]Rancho Los Amigos National Rehabilitation Center, Downey, CA 90242, USA

[3]Department of Neurology, UCI, Irvine, CA 92697, USA

[4]Department of Electrical Engineering and Computer Science, UCI, Irvine, CA 92697, USA

Artifact suppression methods such as adaptive filtering [7], template subtraction [8], and dipole cancellation [9] have been developed to ward against artifacts at analog front-ends. Even if successful, these methods are not fully effective, and so residual artifacts must be suppressed by further processing in the digital back-end. We previously introduced one such signal processing method based on pre-whitening and null projection (PWNP) [10]. It compared favorably to the state-of-the-art methods such as independent component analysis (ICA) while being considerably simpler to implement. Specifically, when tested with electroencephalogram (EEG) data contaminated with sinusoidal artifacts, this method achieved a 30-40 dB artifact suppression [10].

This work seeks to further validate our PWNP method by evaluating its performance on clinically recorded ECoG data. The motivation for our work is two-fold. First, the characteristics of artifact propagation in ECoG appear to significantly differ from those in EEG. Specifically, EEG artifacts arrive at sensors at a variety of phase lags [10], whereas those in ECoG are nearly phase-locked [11]. Second, artifacts caused by the pulse trains used in ECoG somatosensory stimulation [3], [4] may prove to be much more challenging to remove in comparison to the "single-tone" artifacts in our prior study [10]. Namely, these pulse trains produce broadband artifact responses that spectrally overlap with physiologically relevant frequencies, especially those in the motor γ band [11]. Nevertheless, our results with clinically recorded ECoG data indicate that our PWNP artifact suppression method retains the advantage over ICA while being significantly easier to implement. This method, alongside the proper front-end artifact safeguards, thus promises to enable bi-directional operation in BCIs for motor and sensory function restoration.

II. METHODS

A. ECoG Data Collection and Pre-processing

This study was approved by the Institutional Review Board of the University of California, Irvine, and Rancho Los Amigos National Rehabilitation Center. We collected ECoG data at the hospital bedside from an epilepsy patient undergoing a clinical cortical mapping procedure. The patient was implanted with a 4×5 ECoG grid with 10 mm spaced platinum contacts (Integra Life Sciences, Plainsboro NJ) over the left frontal area of the brain (see Fig. 1). Grid implantation and cortical stimulation were performed solely

based on clinical needs. The patient signed an informed consent form to participate in this study.

We recorded ECoG data at a sample rate of $f_s = 512$ Hz. The mapping procedure delivered 5-second-long, biphasic, square pulse trains across a pair of electrodes (stimulating channel) at amplitudes ranging from 2 mA to 10 mA. ECoG signals from the stimulating electrodes were saturated, leaving 18 electrodes for analysis. In the interest of space, this paper shows the analysis of one representative 5-second stimulation event with a 10 mA amplitude, which creates the strongest artifacts. These stimulation-contaminated ECoG data were isolated, along with 5 seconds of baseline ECoG data immediately preceding the stimulation event. Both data sets were high-pass filtered (4$^{\text{th}}$ order, Butterworth, 1.5 Hz).

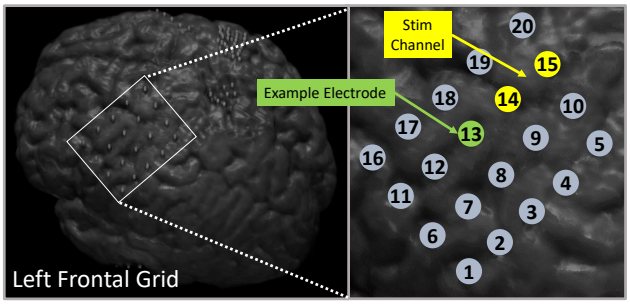

Fig. 1. Left: MR-CT co-registered images showing the placement of the 4 × 5 ECoG grid. Right: Stimulation channel is highlighted yellow and worst-case example electrode is highlighted green on the zoomed-in image.

B. Pre-whitening Null Projection (PWNP) Algorithm

A detailed description of the PWNP algorithm can be found in [10] and its theoretical basis in [12]. The method utilizes the fact that artifacts are much stronger than neural activity and thus reside in a low-dimensional subspace. The signals are then projected away from this subspace by means of its orthogonal complement, so that the projected data reside in an artifact-free subspace. This projection is preceded by a pre-whitening step, which removes spatial correlations across electrodes and improves the signal-to-noise ratio (SNR). These steps can be described by:

$$\mathbf{X}_S^{\text{clean}} = \mathbf{\Sigma}_B^{\frac{1}{2}} \mathbf{H}\mathbf{H}^{\mathrm{T}} \left[\mathbf{\Sigma}_B^{-\frac{1}{2}} (\mathbf{X}_S - \mu_S \mathbf{1}^{\mathrm{T}}) \right] + \mu_S \mathbf{1}^{\mathrm{T}} \quad (1)$$

where $\mathbf{X}_S^{\text{clean}} \in \mathbb{R}^{n \times t_S}$ and $\mathbf{X}_S \in \mathbb{R}^{n \times t_S}$ (n: number of electrodes, t_S: number of time samples) are the "cleaned" and original stimulation-contaminated ECoG data segments, respectively, $\mathbf{\Sigma}_B^{-\frac{1}{2}}$ is the pre-whitening matrix (estimated from the baseline data), μ_S is the time-average of $\mathbf{X}_S$, and $\mathbf{1}$ is a column vector whose elements are 1. $\mathbf{H}$ is the so-called null projection matrix whose columns are the left singular vectors of $\mathbf{\Sigma}_B^{-\frac{1}{2}} (\mathbf{X}_S - \mu_S \mathbf{1}^{\mathrm{T}})$ corresponding to its smallest $n - d$ singular values (d-the dimension of artifact subspace). Theoretically, we showed that these singular values satisfy $\sigma = \sqrt{t_S - 1}$ [10]. To account for noise, it is prudent to identify these singular values by the following criterion: $\sigma \leq \alpha\sqrt{t_S - 1}$, where $\alpha > 1$ [10].

C. ICA Suppression Procedure

To draw a comparison with the PWNP algorithm, an ICA-based suppression procedure was applied to the stimulation-contaminated ECoG data. ICA was performed using the FastICA toolbox in MATLAB [13]. The independent components were individually inspected in the time and frequency domains to identify those primarily consisting of artifacts. These components were then removed, and cleaned data were reconstructed using the back-projection method [14].

D. Artifact Suppression Evaluation

We compared the outputs of the PWNP algorithm and the ICA procedure to determine which method achieved superior artifact removal. The power spectral density (PSD) was calculated for the PWNP-cleaned and ICA-cleaned data, as well as the stimulation-contaminated and baseline data. The integrated power was calculated for each electrode and condition by integrating over frequency and normalizing by the frequency range:

$$P_{\text{int}} = \frac{2}{f_{\text{s}}} \int_0^{\frac{f_{\text{s}}}{2}} PSD(f)\, \mathrm{d}f \quad (2)$$

To quantify the amount of artifact interference introduced to the signal by the stimulation, we calculated the difference between stimulation-contaminated and baseline integrated power. These values also served to establish a baseline for the cleaning procedures. We also quantified the residual artifacts remaining after applying either the PWNP or ICA artifact-removal procedures by calculating the difference between cleaned and baseline integrated power. Finally, we calculated the difference between stimulation-contaminated and cleaned integrated power to determine the removed artifact power for either suppression method. These differences were calculated for each electrode in the grid, interpolated, and mapped onto the co-registered MR-CT images for visualization purposes.

We also calculated the signal-to-interference ratio (SIR) before and after data cleaning at an electrode (LFG13) adjacent to the stimulation channel (Fig. 1). We chose this electrode as a worst-case example, since it is the most significantly impacted by artifacts. To calculate the SIR, the data for each condition were each split up into 5 equal-length segments, and the PSD was calculated for each segment. The mean and standard deviation of power were calculated at each frequency and the SIR was then expressed as an interference index [11]:

$$I(f) = \frac{1}{2} \log \frac{\sigma_{\text{t}}^2(f)}{\sigma_{\text{a}}(f)\sigma_{\text{b}}(f)} \quad (3)$$

where $\sigma_{\text{a}}(f)$ is the standard deviation of either stimulation-contaminated or cleaned PSD at each frequency, $\sigma_{\text{b}}(f)$ is the standard deviation of the baseline PSD, and $\sigma_{\text{t}}(f)$ is the total standard deviation, obtained from the combined stimulation/baseline or cleaned/baseline data. Eq. (3) is a variant of the deflection coefficient [15] that has been modified such that it accounts for overlapping means and unequal variances between the two compared PSDs [16]. In general, it measures the statistical difference between data

distributed over two or more classes. We also performed the Kolmogorov-Smirnov (KS) test (p=0.01) at each frequency to determine whether the PSD for stimulation/cleaned data was significantly different from that of the baseline data.

III. Results and Discussion

Artifacts recorded from the ECoG electrodes appeared as large voltage transients with amplitude ranging from $\sim$100 μV to $\sim$1500 μV (Fig. 3F) and decreasing with distance from the stimulating channel. These transients occurred at the frequency of the pulse train (50 Hz) and were phase-locked within one time sample ($\sim$2 ms) across all electrodes. The five second stimulation duration allowed for approximately 250 artifacts to be captured on each electrode. Fig. 2 shows the PSD of the stimulation-contaminated data at the worst-case electrode. The spectrum exhibited strong peaks at 50 Hz and its super-harmonics (100, 150, 200 Hz), and milder peaks at inter-harmonic frequencies.

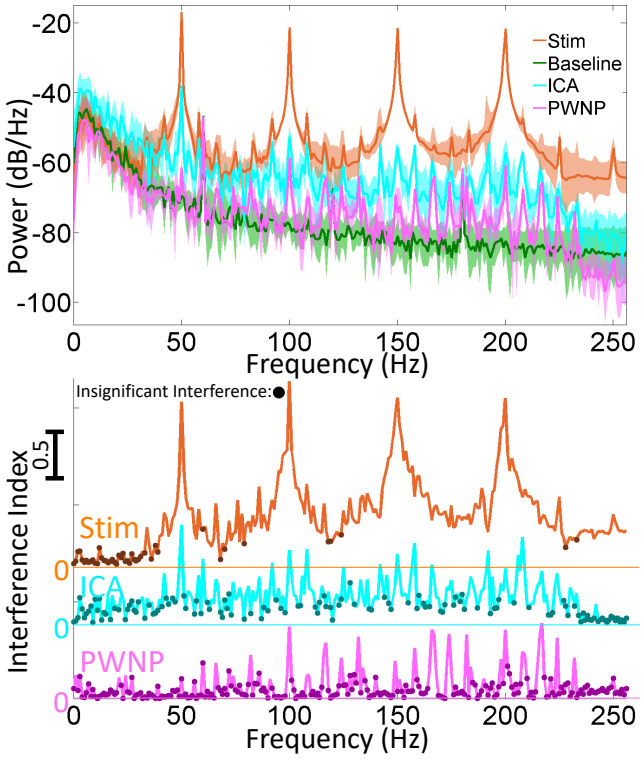

Fig. 2. Top: Average PSD for stimulation-contaminated, baseline, ICA-cleaned, and PWNP-cleaned data for worst-case electrode (LFG13). Shaded area indicates a one standard deviation boundary. Bottom: Interference index, $I(f)$, for stimulation-contaminated, ICA-cleaned, and PWNP-cleaned data. Marked points indicate frequencies that exhibit PSDs that are not significantly different from the PSD at the baseline (KS test, p=0.01).

Since the patient performed no behavioral tasks during baseline or stimulation epochs, we assumed that the cleaned data should take on characteristics similar to the baseline data. In this regard, the PWNP algorithm brought the cleaned data PSD significantly closer to the baseline PSD (Fig. 2). Specifically, the harmonic peaks, broadband power and inter-harmonic peaks were significantly attenuated. This was achieved by choosing $\alpha = 1.1$ (see Section II-B), which caused the PWNP algorithm to identify 11 components (i.e. $d = 11$) residing in the artifact subspace. This is considerably higher than the $d = 3$ estimated in our previous EEG study [10]. However, this discrepancy could be explained by the complexity of the broadband ECoG response elicited by the pulse train stimulation, which required more components to capture. Time-domain inspection of these components showed that they contained voltage transients similar in appearance to the stimulation artifacts. Also, their frequency-domain analysis confirmed the presence of strong spectral peaks at 50 Hz and its super-harmonics. Therefore, by projecting away from this artifact subspace, the transformation (1) substantially suppressed the stimulation artifacts in the data. The difference in integrated power (2) for the worst-case electrode showed that the PWNP algorithm was able to achieve a maximum artifact suppression of 21.49 dB.

The ICA procedure was also able to bring the stimulation-contaminated data PSD closer to that of the baseline data, though not as successfully as the PWNP algorithm (Fig. 2). FastICA was executed on stimulation-contaminated data with default parameters, which resulted in 18 components converging. Upon inspection in the time and frequency domain, 11 were selected manually for removal. In comparison to the PWNP method, the ICA method was only capable of achieving a maximum of 12.01 dB artifact suppression at the worst-case electrode. Though greater attenuation could be achieved by removing more components, inspection of the remaining components did not indicate any that contained significant artifacts. While the subset of removed components could be further optimized, this would require an exhaustive combinatorial search. This is unlike the PWNP algorithm, where the artifact-dominant components are sorted by the singular values and can be readily identified for removal. The ICA procedure thus achieves poorer results with the added burden of having to select artifact components manually or through an exhaustive combinatorial search.

Fig. 2 also shows the interference index across the frequency range for the worst-case electrode. Not surprisingly, the interference index of the stimulation-contaminated data resembled its PSD, with the largest interference occurring at 50 Hz and its super-harmonics. The interference index of both the ICA- and PWNP-cleaned data was significantly reduced, both at the fundamental and super-harmonic frequencies, as well as across the entire frequency range. This was corroborated by the KS test, where the number of frequency points with insignificant interference increased from 45 (Stim) to 125 (ICA) and 219 (PWNP). Overall, these results are consistent with the PWNP method's superior performance at the worst-case electrode.

To quantify the artifact suppression performance across the entire ECoG grid, we created the power difference maps (see Fig. 3). Fig. 3A demonstrates the overall level of artifact contamination, with values reaching as high as 29.18 dB. As expected, the strongest artifacts were seen on electrodes closest to the stimulation channel. Figs. 3B and 3C show the artifact contamination after the data were cleaned with the PWNP and ICA methods, respectively. Both methods

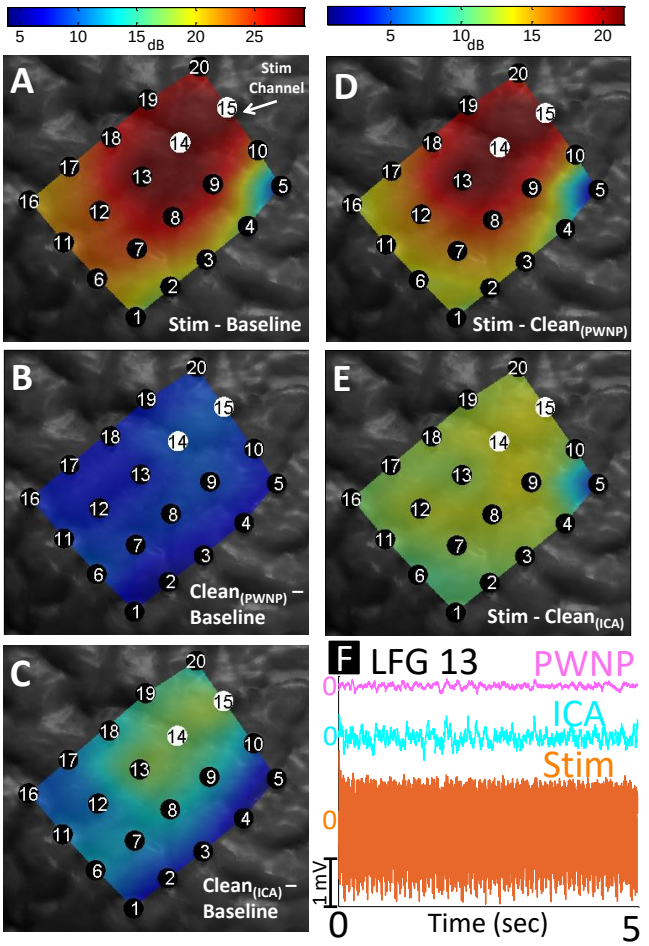

Fig. 3. Power difference maps showing the effect of the PWNP and ICA procedures. Power differences were calculated for each electrode by subtracting the integrated power (Eq. 2) of the second condition from that of the first condition (see Section II-D). These values were interpolated across the grid, color-coded, and super-imposed on the MR-CT co-registered image. Since the signals at the stimulation channels were saturated, we imputed values for these electrodes with highest difference in power for each image. The difference in the integrated power between stimulation-contaminated and baseline data (A), PWNP-cleaned and baseline data (B), ICA-cleaned and baseline data (C), stimulation-contaminated and PWNP-cleaned data (D), and stimulation-contaminated and ICA-cleaned data (E). Fig. 3F shows the time domain before and after the cleaning process.

brought the power level much closer to the baseline, with the values of the PWNP method ranging from 5.0 dB to 9.7 dB, and the values of the ICA method from 4.0 dB to 17.0 dB. Though the ICA method achieved the lowest values, these occurred on the electrodes that were the farthest from the stimulation channel and contained the weakest artifacts. Additionally, ICA failed to achieve the same degree of artifact suppression as the PWNP method on the electrodes closest to the stimulation channel, which were highly artifact-susceptible. An alternative view of these results is given in Figs. 3D and 3E, which show the removed artifact power of the PWNP and ICA methods, respectively. As mentioned previously, the worst-case electrode exhibited 21.49 dB of artifact suppression with the PWNP algorithm, and only 12.01 dB with the ICA procedure.

IV. Conclusions

We demonstrated the ability of the PWNP algorithm to effectively suppress stimulation artifacts in clinically collected human ECoG data. In comparison to more traditional ICA methods, our algorithm achieved a greater degree of suppression. Our future works will focus on further validating the algorithm by analyzing data from additional stimulation events across multiple subjects. If successfully validated, our simpler algorithm could be readily incorporated into a BD-BCI, enabling uninterrupted closed loop operation.

References

[1] W Wang, JL Collinger, AD Degenhart, EC Tyler-Kabara, AB Schwartz, DW Moran, DJ Weber, B Wodlinger, RK Vinjamuri, RC Ashmore, et al. An electrocorticographic brain interface in an individual with tetraplegia. *PloS one*, 8(2):e55344, 2013.

[2] MJ Vansteensel, EGM Pels, MG Bleichner, MP Branco, T Denison, ZV Freudenburg, P Gosselaar, S Leinders, TH Ottens, MA Van Den Boom, et al. Fully implanted brain–computer interface in a locked-in patient with ALS. *New Engl. J. Med.*, 375(21):2060–2066, 2016.

[3] SV Hiremath, EC Tyler-Kabara, JJ Wheeler, DW Moran, RA Gaunt, JL Collinger, ST Foldes, DJ Weber, W Chen, ML Boninger, et al. Human perception of electrical stimulation on the surface of somatosensory cortex. *PloS one*, 12(5):e0176020, 2017.

[4] B Lee, D Kramer, MA Salas, S Kellis, D Brown, T Dobreva, C Klaes, C Heck, C Liu, and RA Andersen. Engineering artificial somatosensation through cortical stimulation in humans. *Frontiers in systems neuroscience*, 12, 2018.

[5] FD Szymanski, M Semprini, FA Mussa-Ivaldi, L Fadiga, S Panzeri, and A Vato. Dynamic brain-machine interface: A novel paradigm for bidirectional interaction between brains and dynamical systems. In *Proc. 33rd Ann. Int. Conf. IEEE EMBS*, pages 4592–4595, 2011.

[6] S Flesher, J Downey, JL Collinger, S Foldes, J Weiss, E Tyler-Kabara, S Bensmaia, A Schwartz, M Boninger, and R Gaunt. Intracortical microstimulation as a feedback source for brain-computer interface users. In *Brain-Computer Interface Research*, pages 43–54. Springer, 2017.

[7] AE Mendrela, J Cho, JA Fredenburg, V Nagaraj, TI Netoff, MP Flynn, and E Yoon. A bidirectional neural interface circuit with active stimulation artifact cancellation and cross-channel common-mode noise suppression. *IEEE J. Solid-St. Circ.*, 51(4):955–965, 2016.

[8] WA Smith, JP Uehlin, SI Perlmutter, JC Rudell, and VS Sathe. A scalable, highly-multiplexed delta-encoded digital feedback ecog recording amplifier with common and differential-mode artifact suppression. In *2017 Symp. VLSI Circ.*, pages C172–C173, 2017.

[9] J Lim, PT Wang, H Pu, CY Liu, S Kellis, RA Andersen, P Heydari, AH Do, and Z Nenadic. Dipole cancellation as an artifact suppression technique in simultaneous electrocorticography stimulation and recording. In *Proc. 9th Int. IEEE/EMBS Conf. Neural Eng.*, pages 725–729, 2019.

[10] PT Wang, CM McCrimmon, P Heydari, AH Do, and Z Nenadic. Subspace-based suppression of cortical stimulation artifacts. In *Proc. 40th Ann. Int. Conf. IEEE EMBS*, pages 2426–2429, 2018.

[11] J Lim, PT Wang, A Bidhendi, O Arasteh, S Shaw, M Armacost, H Gong, C Liu, A Do, P Heydari, and Z Nenadic. Characterization of stimulation artifact behavior in simultaneous electrocorticography grid stimulation and recording. In *Proc. 40th Ann. Int. Conf. IEEE EMBS*, pages 4748–4751, 2018.

[12] S-C Wu, AL Swindlehurst, PT Wang, and Z Nenadic. Projection vs. prewhitening for EEG interference suppression. *IEEE T. Bio-med. Eng.*, 59(5):1329–1338, 2012.

[13] A Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE T. Neural Networ.*, 10(3):626–634, 1999.

[14] F Cong, T Ristaniemi, and H Lyytinen. Advanced signal processing on event-related potentials (erps).

[15] SM Kaye. Fundamentals of statistical signal processing. detection theory, 1989.

[16] Z Nenadic. Information discriminant analysis: Feature extraction with an information-theoretic objective. *IEEE T. Pattern Anal.*, 29(8):1394–1407, 2007.