Series: Smart Science Design and Technology 1

ENGINEERING INNOVATION AND DESIGN

editors
Artde D.K.T. Lam
Stephen D. Prior
Siu-Tsen Shen
Sheng-Joue Young
Liang-Wen Ji

CRC Press
Taylor & Francis Group

ENGINEERING INNOVATION AND DESIGN

Smart Science, Design and Technology

ISSN: 2640-5504
eISSN: 2640-5512

Book series editors

Stephen D. Prior
Faculty of Engineering and Physical Sciences, University of Southampton, Southampton, UK

Siu-Tsen Shen
Department of Multimedia Design, National Formosa University, Taiwan, R.O.C.

Volume 1

PROCEEDINGS OF THE 7TH INTERNATIONAL CONFERENCE ON INNOVATION, COMMUNICATION AND ENGINEERING (ICICE 2018), HANGZHOU, CHINA, 9–14 NOVEMBER 2018

Engineering Innovation and Design

Editors

Artde D.K.T. Lam
Fujian University of Technology, P.R. China

Stephen D. Prior
University of Southampton, UK

Siu-Tsen Shen, Sheng-Joue Young & Liang-Wen Ji
National Formosa University, Yunlin, Taiwan

CRC Press
Taylor & Francis Group
Boca Raton London New York Leiden

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A BALKEMA BOOK

CRC Press/Balkema is an imprint of the Taylor & Francis Group, an informa business

Typeset by V Publishing Solutions Pvt Ltd., Chennai, India

Published by: CRC Press/Balkema
Schipholweg 107C, 2316 XC Leiden, The Netherlands
e-mail: Pub.NL@taylorandfrancis.com
www.crcpress.com – www.taylorandfrancis.com

ISBN: 978-0-367-02959-3 (Hbk)
ISBN: 978-0-429-01977-7 (eBook)
DOI: https://doi.org/10.1201/9780429019777

Engineering Innovation and Design – Lam et al. (Eds)
© 2019 Taylor & Francis Group, London, ISBN 978-0-367-02959-3

Table of contents

Engineering Innovation and Design – Lam et al. (Eds)
© 2019 Taylor & Francis Group, London, ISBN 978-0-367-02959-3

Editorial board

Advanced material science and engineering

Engineering Innovation and Design – Lam et al. (Eds)
© 2019 Taylor & Francis Group, London, ISBN 978-0-367-02959-3

Effect of temperature on structure and properties of MoS_2 crystal grown by chemical vapor deposition

Jiaying Jian, Honglong Chang, Jiaojie Yue & Pengfan Dong
School of Electronic and Information Engineering, Xi'an Technological University, Xi'an, P.R. China

ABSTRACT: To meet the need for preparing high-performance two-dimension materials electrical and optical devices based on single-layer MoS_2, the effects of the heating temperature of S and MoO_3 on the morphology, size, structure and layers of a MoS_2 crystal grown on sapphire substrate by using chemical vapor deposition are studied in this paper. The results show that the growth of MoS_2 crystal depends on the temperature. Either a too low or too high of the heating temperatures of S and MoO_3 are not conducive to the formation of MoS_2. When the temperature of S is in the range from 160 to 180°C and the temperature of MoO_3 is in the range from 795 to 820°C, the size of MoS_2 increases with increasing the temperature. A uniform large-sized triangle with a side length of 90 μm is obtained when the heating temperature of MoO_3 is 820°C and the heating temperature of S is 180°C.

1 INTRODUCTION

Two-dimensional (2D) crystals of transition metal dichalcogenides (TMDs) have been widely studied because of their excellent electrical and optical properties. (Wang, et al., 2011; Jariwala, et al., 2014; Cunningham, et al., 2012) Molybdenum disulfide (MoS_2) is a prototypical TMD. Bulk MoS_2 is a semiconductor with an indirect bandgap of 1.2 eV. (Li, et al., 2012) When the thickness of a MoS_2 film reaches the single-layer level, the bandgap will eventually increase by at least 0.6 eV due to the two-dimensional confinement effect. The reduction in thickness eventually causes the bandgap of the single-layer MoS_2 film to change from an indirect bandgap to a direct one. This change in bandgap results in the photoluminescence of high brightness in single-layer MoS_2 crystal. (Splendiani, et al., 2010) In addition, single-layer MoS_2 crystal has prominent electrical performance, and the transport calculation of the nonequilibrium Green's function shows that the switching ratio of MoS_2 can reach up to 10^{10}. (Yoon, et al., 2011) Due to its superior electrical and optical performance, the single-layer MoS_2 crystal has a great potential application prospect in the electronic devices and photoelectronics fields. (Mak, et al., 2010; Lopezsanchez, et al., 2013; Mak, et al., 2011; Radisavljevic, et al., 2013).

The common methods for preparing single-layer MoS_2 include micromechanical exfoliation, ultrasonic-assisted liquid-phase exfoliation, intercalation of lithium ions, and the chemical vapor deposition (CVD) method. (Ferraz, et al., 2016; Xia, et al., 2017; Huang, et al., 2018; Lee, et al., 2012) The micromechanical exfoliation method is simple to operate, and the cleavage products are mostly single-layer MoS_2 with high purity and carrier mobility. Its disadvantage is that it is hard to control the number and size of the cleavage MoS_2 layers. The ultrasonic-assisted liquid-phase exfoliation method can prepare single-layer MoS_2 nanosheets. Moreover, the obtained sheet can be easily assembled into a film. Its disadvantage is that the ultrasonic power has a great influence on the formation of the nanosheet. The intercalation of lithium ions method can prepare high-quality single-layer MoS_2 materials through controlling the lithium ion insertion and exfoliation process by an electrochemical lithium battery device. This method has the advantages of large size and high quality. But, its operation process is complicated and exfoliation efficiency is very low.

The chemical vapor deposition (CVD) method uses sulfur vapor to chemically react with molybdenum oxide vapor to form MoS_2, (Amani, et al., 2016; Liu, et al., 2013) which can be applied to manufacture devices such as new-type resonators and transistors based on two-dimensional materials. (Yin, et al., 2012; Lee, et al., 2013) The single-layer MoS_2 prepared by this method has excellent optical and electrical performance. The structure and morphology of sediment on the substrate depend on vapor pressures of S and MoO_3, as well as the reaction temperature and time, which depend on the heating temperatures of S and MoO_3 powders, and the gas flow rate. Single-layer MoS_2 films with a size of 20 to 80 μm have been grown by CVD.

However, whether larger-sized single-layer MoS_2 films can be prepared by CVD remains unsolved.

In this paper, the effects of the heating temperatures of S and MoO_3 powders on the morphology, the size, the structure and layer number of MoS_2 crystal grown on a sapphire substrate by sulfurized MoO_3 via the CVD method.

2 EXPERIMENTAL

2.1 *Growth of the MoS_2 crystal*

Sapphire plate was chosen as the substrate. First of all, the sapphire substrate was cleaned in a clean room. The sapphire substrate was first sonicated in acetone, ethanol and deionized water for 10 min, respectively, and then dried with a nitrogen gas gun.

A chemical vapor deposition double temperature tube furnace, as shown in Fig. 1, was adopted to grow MoS_2 crystal by vapor deposition. The S powder was first placed in the low temperature zone, the MoO_3 powder was placed in the high temperature zone, and the sapphire substrate was placed flat on the right side of the furnace tube. The sulfur source in the low temperature zone was 15 cm away from the molybdenum source in the high temperature zone, and the molybdenum source was 5 cm away from the sapphire substrate. The purities of the S powder and MoO_3 powder are 99.5%.

After the tube furnace was vacuumed to 10 Torr with a mechanical pump, it was filled with high-purity argon gas at a flow rate of 70 sccm for 30 min. The evacuation-fill processes were repeated three times to vent the air and impurities in the tube furnace. After that, high-purity argon gas was introduced into the tube furnace, and when the pressure reached 1.2 atm, the exhaust valve was opened and adjusted to make the pressure in the furnace greater than 1 atm. Then, the furnace was heated by using a two-step heating method, as shown in Fig. 2. First, the MoO_3 zone was rapidly heated to 600°C at a rate of 20°C/min and then slowly heated to the setting temperature (770°C, 795°C, 820°C, 845°C) at a rate of 5°C/min. The heating of the S zone began at the time when the temperature of the MoO_3 zone was heated to 600°C.

2.2 *Characterization of the MoS_2 crystal*

A scanning electron microscope (SEM, FEI-QUANTA-400) was used to observe the morphology of the crystal sample grown by CVD.

The layer number of MoS_2 was determined by Raman spectroscopy (Renishaw-UK, with a light-

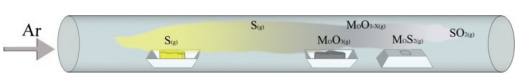

Figure 1. Schematic of growth MoS_2 crystal by CVD.

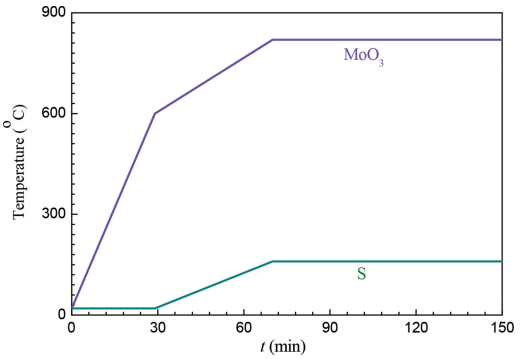

Figure 2. The heating and holding temperature profile of MoO_3 source and S source in double temperature tube furnace.

passing efficiency of more than 30%, a spectral range of 200 to 1000 nm, a spectral resolution of 1 cm^{-1}, and a spatial resolution of 0.5 μm in the lateral direction and 2 μm in the longitudinal direction) and a homemade photoluminescence spectrometer.

Based on the Raman scattering effect, Raman spectroscopy can be used to study the structure of molecules by analyzing the scattering spectrum different from the incident light frequency. There are five vibration modes (E^2_{2g}, E^1_{2g}, E_{1g}, E_{1u} and A_{1g}) in MoS_2. For the Raman spectroscopy test of MoS_2, only E^1_{2g} and A_{1g} modes are observed. The characteristic peak position of the Raman spectrum of MoS_2 is related to the film thickness. The Van der Waals force increases as the layer number reduces. When the layer number of the film reduces gradually, the characteristic peak corresponding to the A_{1g} vibration mode happens to red shift (shifts to the long wavelength side), and the characteristic peak corresponding to the E^1_{2g} vibration mode happens to blue shift (shifts towards the blue end of the spectrum). When MoS_2 is a bulk material, the E^1_{2g} peak is located near 382 cm^{-1}, and the A_{1g} peak is located near 407 cm^{-1}. When the layer number of MoS_2 is reduced to a single layer, the E^1_{2g} peak shifts towards the blue end to approximately 385 cm^{-1}, and the A_{1g} peak shifts to the long wavelength side to approximately 403 cm^{-1}. The A_{1g} peak corresponding to MoS_2 film of different layers differs from the E^1_{2g} peak, and thus the layer number of MoS_2 could be determined based on the difference between the detected A_{1g} peak and E^1_{2g} peak.

The fluorescence effect of the bulk structure of MoS_2 semiconductor material is not obvious. When it is gradually thinned to a few layers or even a single layer, the band gap of MoS_2 changes to a direct band gap, and the fluorescence effect is obviously enhanced, with two characteristic peaks appearing near 1.8 eV and 2.0 eV. The photoluminescence

spectrum can be used as an effective means to characterize the layer number of MoS_2.

3 RESULTS AND DISCUSSION

3.1 *Effect of temperatures on the morphology of MoS_2 crystal*

(1) S temperature: The deposition of MoS_2 crystal by S vapor and MoO_3 vapor involves Reactions (1)–(3).

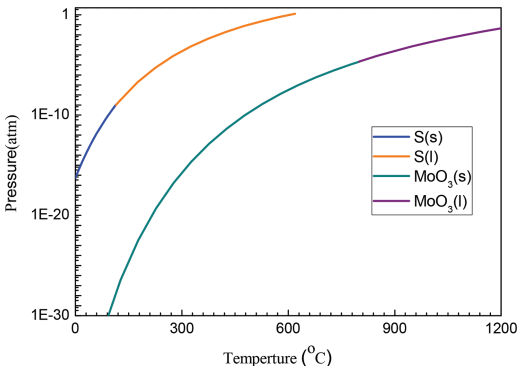

Figure 3. Dependence of the vapor pressure P on the temperature T for MoO_3 and S.

$$MoO_{3(g)} + (x/2)S_{2(g)} = MoO_{3-x(g)} + (x/2)SO_2 \quad (1)$$

$$MoO_{3-x(g)} + (7-x)/S_{2(g)} = MoS_{2(g)} + (3-x)/2SO_2 \quad (2)$$

$$MoS_{2(g)} = MoS_{2(s)} \quad (3)$$

The vapor pressures of MoO_3 and S and the temperature are the decisive factors affecting the deposition rate of MoS_2 crystal. The dependence curves of the vapor pressure on the temperature for MoO_3 and S is shown in Fig. 3. It can be seen that the vapor pressures of MoO_3 and S increase with increasing the temperature. Therefore, increasing the vapor pressures of MoO_3 and S can promote the formation of MoS_2 crystal.

First of all, the effect of S vapor pressure (the temperature of S zone) on the formation of monolayer MoS_2 crystals is studied under other (molybdenum source dosage, molybdenum temperature, holding time, gas flow rate, distance between the carrier and the molybdenum) fixed condition. Fig. 4 shows the effect of S temperature (140, 160, 180 and 200°C) on the morphology of MoS_2 crystal obtained by SEM.

It can be seen that, as shown in Fig. 4(a), a mixture of white MoO_2 particles and rhombic $MoS_{2(s)}$ crystals is obtained when the temperature of S is 140°C. The reason for this is that the vapor pressure of S is as low as $10^{-7.96}$ atm, which makes

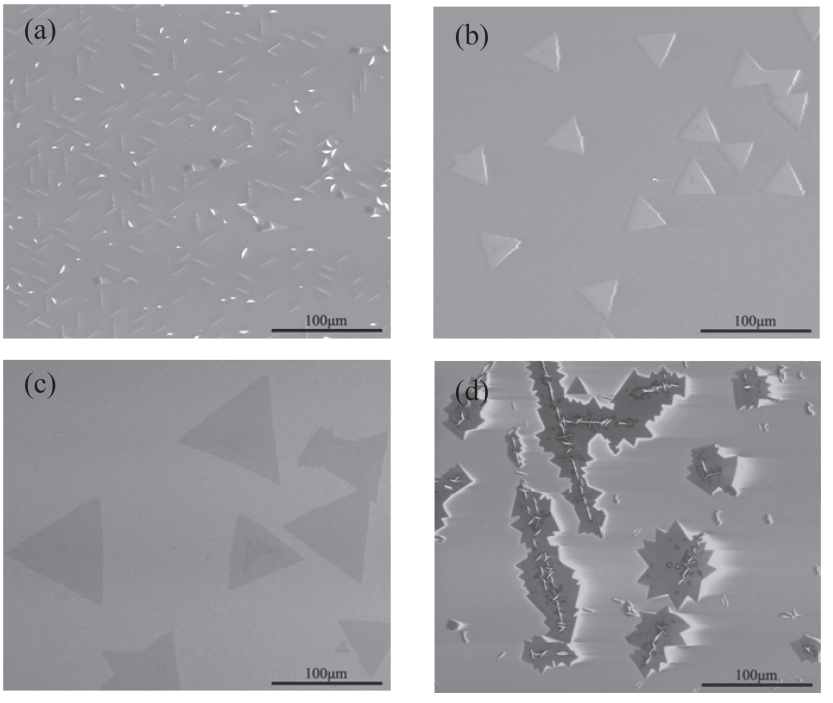

Figure 4. SEM images of MoS_2 prepared from different S temperatures. (a) 140°C; (b) 160°C; (c) 180°C; (d) 200°C.

$MoO_{3-x(g)}$ cannot be completely reduced by S in time. The excessive $MoO_{3-x(g)}$ can decompose into $MoO_{2(s)}$ at high temperature, as shown in Reaction (4). So the deposits on the substrate are a mixture of $MoO_{2(s)}$ and $MoS_{2(s)}$.

$$MoO_{3-x(g)} = MoO_{2(s)} + (1/2 - x/2)O_2 \quad (3)$$

When the temperature of S increases from 140 to 160°C, as shown in Fig. 3, its vapor pressure increases five folds. This can make more S to react with MoO_3 and eliminate the excess of $MoO_{3-x(g)}$, which avoid the formation of $MoO_{2(s)}$. So, as shown in Fig. 4(b), there are no MoO_2 white particles on the substrate and the surface is very clean. The obtained MoS_2 crystal is the triangular one with a size of about 30 μm.

When the temperature of S is 180°C, as shown in Fig. 4(c), an amount of complete and large-size triangular MoS_2 crystals with a side length of approximately 90 μm appears. This result from that increasing the temperature of S can increase its vapor pressure, providing more S to react with MoO_3. Thus, larger size of MoS_2 is formed.

When the temperature of S is higher than 200°C, the MoS_2 crystal number increases quite a bit. But the crystal size does not become larger. As shown in Fig. 4(d), a large number of regular triangular MoS_2 crystals are contacted or stacked together to form multi-layer MoS_2. The critical size of the atomic cluster for MoS_2 to form nucleation decreases with increasing concentration of $S_{2(g)}$. The decrease of the critical size of the atomic cluster for MoS_2 to form nucleation can lead to the continuous formation of MoS_2 crystal nucleus, thus suppressing its growth. Therefore, when the temperature of S is higher than 200°C, a large number of multi-layer MoS_2 crystals are formed.

(2) MoO_3 temperature: The effect of MoO_3 vapor pressure on the preparation of monolayer MoS_2 crystals is studied under the condition of constant vapor pressure of $S_{(g)}$ at 180°C. Fig. 5 shows the effect of MoO_3 temperature (770°C, 795°C, 820°C, and 845°C) on the morphology of single-layer MoS_2 crystal obtained by SEM. It can be seen that, as shown in Fig. 5(a), there is almost nothing on the substrate when the temperature of MoO_3 is 770°C. It results from that the amount of MoO_3 participating in the chemical vapor deposition reaction is insufficient since there is no enough MoO_3 sublimated at 770°C (the evaporation pressure of MoO_3 is only $10^{-5.02}$ atm).

The evaporation pressure of MoO_3 increase with increasing temperature. As shown in Fig. 3, when the temperature of MoO_3 is 795°C, the evaporation pressure of MoO_3 is 2 times compared with that at

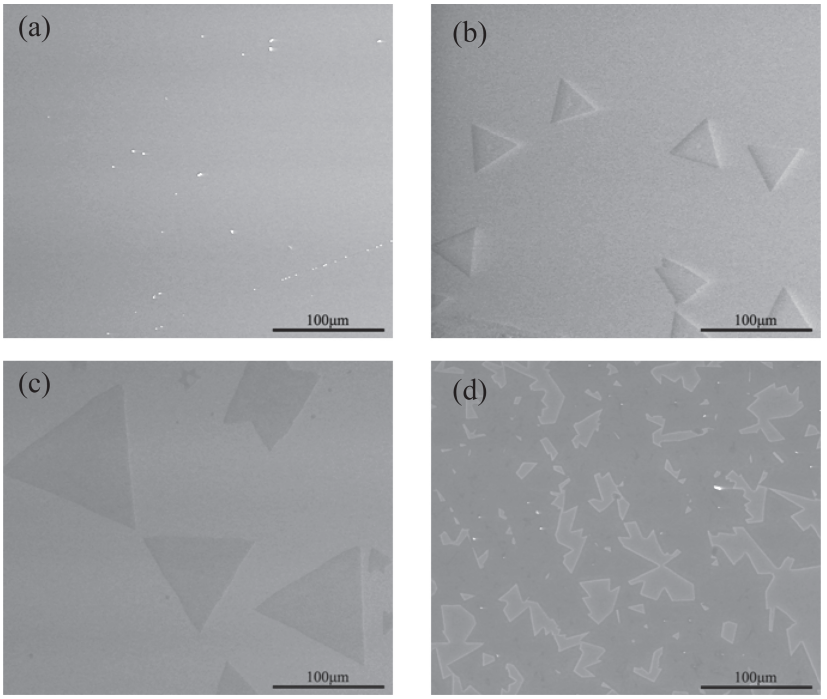

Figure 5. SEM images of MoS_2 prepared from different MoO_3 temperatures. (a) 770°C; (b) 795°C; (c) 820°C; (d) 845°C.

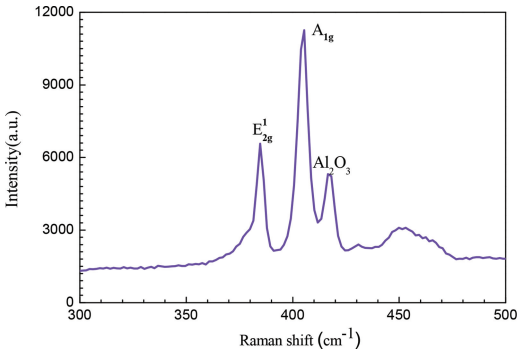

Figure 6. Raman spectra of MoS_2.

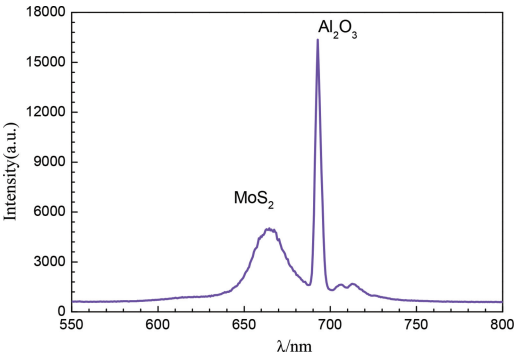

Figure 7. Photoluminescence spectra of MoS_2.

770°C, as shown in Fig. 5(b). This makes the amount of MoO_3 participating in the chemical vapor deposition reaction be sufficient. Therefore, triangular MoS_2 with the side length of 30 μm is obtained. Obviously, the size of the triangular single-layer MoS_2 on the substrate is not sufficiently large.

The size of the single-layer MoS_2 depends on the evaporation pressure of MoO_3. When the temperature of MoO_3 is 820°C, the evaporation pressure of MoO_3 is 3.9 times compared with that at 770°C. As shown in Fig. 5(c), an amount of complete and independent triangular MoS_2 crystals with a side length of approximately 90 μm appears. This results from the higher temperature increasing the evaporation pressure of MoO_3 and the amount of MoO_3 evaporated, providing sufficient $MoO_{3(g)}$ to react with $S_{2(g)}$. So large-size MoS_2 is obtained when the temperature of MoO_3 is 820°C.

When the temperature of MoO_3 is 845°C, as shown in Fig. 5(d), the size of MoS_2 does not become larger but decreases. However, the MoS_2 crystal number increases in this condition. A large number of regular triangular MoS_2 crystals are contacted or stacked together to form multi-layer MoS_2. When the temperature of MoO_3 is increased to 845°C, the evaporation pressure of MoO_3 ($10^{-4.16}$ atm) increases 7.3 times compared with that at 770°C. This can increase the concentration of $MoS_{2(g)}$ in the furnace. The critical size of the atomic cluster for MoS_2 to form nucleation decreases with increasing concentration of $MoS_{2(g)}$. The decrease of the critical size of the atomic cluster for MoS_2 to form nucleation can lead to the continuous formation of MoS_2 crystal nucleus, thus suppressing its growth. Therefore, when the temperature of MoO_3 is 845°C, a large number of multi-layer MoS_2 crystals are formed.

In addition, at this higher temperature, the evaporation rate of MoO_3 is much higher, and $MoO_{3(g)}$ cannot be completely reduced by $S_{2(g)}$ in time. The decomposition of the excessive $MoO_{3(g)}$ can form $MoO_{2(s)}$. Therefore, when the temperature of MoO_3 is 845°C, the deposits on the substrate are a mixture of $MoO_{2(s)}$ and $MoS_{2(s)}$.

3.2 *Characterization of MoS_2 structure*

(1) Characterization by Raman Spectroscopy: Fig. 6 shows the results of single-layer triangular MoS_2 samples obtained by Raman spectroscopy, in which the temperature in the high temperature zone is 820°C and the temperature in the low temperature zone of the sample is 180°C of the sample in Fig. 5(c). It can be seen that the E^1_{2g} peak of the single-layer triangular MoS_2 in the two samples is approximately 385 cm^{-1}, and the A_{1g} peak is approximately 405 cm^{-1}. The wave number difference is around 20 cm^{-1}. According to the literature, the peak appearing at a wave number of 416 cm^{-1} is caused by the influence of the sapphire substrate Al_2O_3. This indicates that the deposits on the sapphire substrate are single-layer MoS_2. This is consistent with the predicted results of SEM.

(2) Characterization of Photoluminescence Spectra: Fig. 7 shows the photoluminescence spectra of two single-layer triangular MoS_2 samples. The sample in Fig. 5(c) shows a strong characteristic peak at the wavelength of λ = 664.7 nm. Through the relationship between the wavelength and electronic Ford conversion, it can be seen that the corresponding transition energy level E = 1.86 eV at the luminescence peak λ = 664.7 nm, which is in accordance with the reported results of the literature, further verifying that the MoS_2 crystals deposited on the substrate in Fig. 8 are of single-layer structure.

4 CONCLUSIONS

(1) The growth of MoS_2 crystal depends on the temperature. Either a too low or too high of the heating temperatures of S and MoO_3 are not

conducive to the formation of MoS_2. When the temperature of S is in the range from 160°C to 180°C and the temperature of MoO_3 is in the range from 795 to 820°C, the size of MoS_2 increases with increasing the temperature. A uniform large-sized triangle with a side length of 90 μm is obtained when the heating temperature of MoO_3 is 820°C and the heating temperature of S is 180°C.

(2) The test results of SEM, Raman spectroscopy and photoluminescence spectroscopy show that the triangular MoS_2 crystals are of single-layer structure when the heating temperature of MoO_3 is 820°C and the heating temperature of S is 180°C.

REFERENCES

Amani, M., R.A. Burke, X. Ji, et al, 2016. *Acs Nano.* 10(7) 6535–6541.

Cunningham, G., M. Lotya, C.S. Cucinotta, et al, 2012. *Acs Nano.* 6(4) 3468–3480.

Ferraz, D.C.M.C., H.B. Ribeiro, F. Kessler, et al., 2016. *Materials Research Express* 3(2) 025303.

Huang, J., X. Deng, H. Wan, et al, 2018. *Acs Sustainable Chemistry & Engineering* 6(4) 157–163.

Jariwala, D., V.K. Sangwan, L.J. Lauhon, et al., 2014. *Acs Nano.* 8(2) 1102–1120.

Lee, J., Z. Wang, K. He, et al., 2013. *Acs Nano.* 7(7) 6086–6091.

Lee, Y.H., X.Q. Zhang, W. Zhang, et al, 2012. *Advanced Materials* 24(17) 2320–2325.

Li, H., Q. Zhang, C.C.R. Yap, et al., 2012. *Advanced Functional Materials* 22(7) 1385–1390.

Liu, K.K., Y. H. Lee, Y.C. Lin, et al, 2013. *Ecs Transactions.* 50(6) 61–63.

Lopezsanchez, O., D. Lembke, M. Kayci, et al., 2013. *Nature Nanotechnology* 8(7) 497.

Mak, K.F., C. Lee, J. Hone, et al., 2010. *Physical Review Letters.* 105(13) 136805.

Mak, K.F., K. He, J. Shan, et al., 2012. *Nature Nanotechnology* 7(8) 494–498.

Radisavljevic, B., M.B. Whitwick, A. Kis, 2013. *Acs Nano.* 7(4) 9934–9938.

Splendiani, A., L. Sun, Y. Zhang, et al., 2010. *Nano Letters.* 10 (4) 1271–1275.

Wang, Q.H., K. Kalantarzadeh, A. Kis, et al., 2012. *Nature Nanotechnology* 7(11) 699–712.

Xia, J., J. Wang, D. Chao, et al, 2017. *Nanoscale*. 9(22) 7533.

Yin, Z., H. Li, H. Li, et al, 2012. *Acs Nano.* 6(1) 74–80.

Yoon, Y., K. Ganapathi, S. Salahuddin, 2011. *Nano Letters.* 11 (9) 3768–3773.

Engineering Innovation and Design – Lam et al. (Eds)
© 2019 Taylor & Francis Group, London, ISBN 978-0-367-02959-3

Influence of alcoholic solvents on the anti-glare property of silica sol-gel thin films

Yu-Hui Huang & Huann-Ming Chou
Department of Mechanical and Energy Engineering, Kun Shan University, Yong-kang, Tainan, Taiwan

Lung-Chuan Chen
Department of Materials Engineering, Kun Shan University, Yong-kang, Tainan, Taiwan

ABSTRACT: This study employed Tetraethoxysilane (TEOS) and Methyltrimethoxysilane (MTMS) in alcohol-water media catalyzed by HNO_3 to synthesize silica sol-gel, wherein different alcohols (methanol, ethanol, propanol, isopropanol, and n-butanol) were conducted. The alcohol types have been demonstrated to exert much impact on the property of prepared silica sol-gel. The particle size of silica sol-gel decreases in order of methanol > ethanol > propanol > isopropanol > n-butanol. The silica sol-gel samples prepared by different solvents were deposited on 3.0 mm soda-lime glass substrates to generate anti-glare silica sol-gel thin films using an auto-spray system. The properties of gloss, haze, transmittance, and surface roughness of the antiglare silica sol-gel films were examined. The results indicate that the average haze of silica sol-gel film/glass substrate descends in the following order of methanol > ethanol > isopropanol > propanol > n-butanol, while the average gloss is isopropanol > propanol > n-butanol > ethanol > methanol. The relative magnitude of average transmittance is methanol > propanol ~ ethanol > n-butanol > isopropanol, but their deviation is very insignificant. In summary, the solvent of methanol shows the highest average transmittance of 92.35%, maximum average haze value of 9.76%, and the lowest average gloss value of 74.3 GU. In contrast, the solvent of isopropanol exhibits the minimum arithmetical mean deviation (Ra) of surface roughness (0.064 um). This study explores the fundamental roles of alcohol solvents which can be used as the basis for subsequent variable optimization of automated spray systems in the future.

1 INTRODUCTION

With the advancement of technology, humans receive a lot of information through electronic products. Electronic displays deliver these messages through human's eyes to our brains. Visual comfort and fine legibility are the basics of the display. The ambient illumination is an important factor affecting visual performance and visual fatigue. Excess illumination may cause annoyance, discomfort, and impair vision due to exceeding brightness of the acceptable level of human's eyes, which is called as glare. Glare originates from direct illumination and reflected illumination from the bodies within the vision field. Glare reduces visual performance, legibility and injuries eyes. Anti-Glare treatment could reduce the high light intensity and glare caused by excessive concentration of light, thereby improving the user's comfort for cover lens. Depositing a rough thin film and laminating a matte surface layer are broadly adopted for antiglare treatment. (Lin, 2008).

Surface atomization process is widely applied to prepare anti-glare glass, and the methods of nanoparticle coating (Ma, et al., 2010), solution etching (Chen, et al., 2016) and sol-gel coating (Brinker, et al., 1990; Klein, 1994) are commonly involved to generate antiglare structures. Although nano-particles can be easily coated on glass substrates by spray coating, spin coating, and dip coating techniques, however, there are still some technical challenges in the uniformity of nanoparticle distribution.

Although it has been over 30 years that the spraycoating process with the sol-gel particles as the industry's main method for anti-glare surface treatment, however, till now, the optimization conditions of these processes are still far from clear because they are affected by more than 50 main variables and their interactions (Lugscheider, et al., 1996). In this report, different solvents are used to synthesize silica sol-gel particles, which are then deposited on soda-lime glass substrates to act as anti-glare coating materials. The particle size of the silica sol-gel and the physical properties of the anti-glare glasses were examined by the surface glossiness, haze, average transmittance, and surface roughness minimum arithmetical mean

deviation (Ra) after the spray-drying process. This is an important pilot study of the optimization of automated spray systems in the future.

2 EXPERIMENTAL

2.1 *Materials*

High purity tetraethoxysilane (TEOS) and methyltrimethoxysilane (MTMS) were purchased from Evonik (Germany). Nitric acid (EP) was purchased from Union Chemical Works. Methanol, n-propanol, and n-butanol were all of 99.9% purity and obtained from Shiny Chemical Industry Co. The purity of ethanol and isopropanol were above 99.9% and supplied by Champion Benefit Enterprises and LCY Chemicals, respectively. Purified Water (> 18 MΩ cm) was used in this study.

2.2 *Fabrication of silica Sol-gel*

TEOS, MTMS, alcohol, water and nitric acid were added in a 1L glass container and then magnetically agitated for 24 h at 25°C. Then, the sol-gel was aged at 4°C for 4 days. The molar ratio of TEOS:MTMS:alcohol:HNO3:H2O was 1:0.39:8.39:0.02:5.17.

2.3 *Preparation of Anti-glare film layer*

The glass substrates with a diameter of 100 × 100 mm and thickness of 3 mm were ultrasonically cleaned at 50°C for 30 min and then baked at 80°C for 1h. The prepared silica sol-gel samples were deposited onto the cleaned substrates through an automated spray coating system. The sol-gel to be delivered pressure was 300 KPa, the provided pressure of transport air was 400 KPa, and the spray gun displacement speed was 300mm/sec under one-pass spray operation conditions. The obtained silica antiglare film/glass samples were heated at 180°C for 1h. In this work, the Auto-spray gun (S-710AD) was purchased from Guan Piin Painting Technology Co., Ltd. (Taiwan). The environmental temperature/humidity was well controlled at 25°C/40%.

2.4 *Characterization*

The effective diameters of the silica sol-gel particles obtained from different alcohol solvents (methanol, ethanol, propanol, isopropanol, nbutanol) were measured using the Nano ZS particle size and zeta potentials analyzer (Malvern Panalytical, Malvern, UK). The surface anti-glare property of silica sol-gel thin films were measured using the BYK micro-TRI-gloss device (BYK Additives & Instruments Company, Bavarian, GER), WGT-S haze meter (Labthink Instruments Company, Jinan, CN) and HMT MFS-630 angleadjustable optical measurement analyzer (Hong-Ming Technology Company, New Taipei City, TW).

3 RESULTS AND DISCUSSION

3.1 *Particle size of silica Sol-gel*

The particle size's distribution of silica sol-gel obtained from different alcohol solvents with keeping other operating variables and reactants' composition constant is depicted in Fig. 1, and Fig. 2 displays the corresponding average particle sizes, which indicates that the particle sizes obtained from methanol, ethanol, propanol, isopropanol, and n-butanol are 4.17 nm, 2.92 nm, 2.07 nm, 1.72 nm, and 1.59 nm, respectively. The results showed that the particle size of silica sol-gel become smaller as the carbon number in alcohol molecules increases. In contrast, the particle size is in coincidence with the dielectric constant of the alcohols. The dielectric constant of methanol, ethanol, propanol, isopropanol, and n-nutanol are 33.7, 25.0, 20.3, 18.7, and 17.8.

With the acidic catalyst in aqueous system, ethoxide groups are highly accessible to be protontaed firstly, and then withdraw electrons form silicon, resulting in the electrophilic hydrolysis reaction with water. It is suggested that the particle size of silica sol-gel primarily depend on the condesnation rather than hydrolysis due to the fast reaction rate of the later. It is likely that alcohol solvent may surround the alkoxide or even react with alkoxide, and this may develop the discrepancy of particle size in silica sol-gel because of satiric effect introduced by alcohols during condensation stages. Furthermore, the literature (Chou, et al., 2003; Matsuyama, et al., 1991; Malay, et al., 2013) indicated that under the acidic condition, the hydrolysis is the main reaction, and the condensation reaction is less pronounced relatively, therefore the particle size was considered to be affected significantly by the condensation reaction.

3.2 *Gloss, haze, transmittance and ra*

The silica sol-gel samples prepared from various alcohols were transferred to the auto-spray systems to deposit an anti-glare layer on 3.0 mm soda-lime glass substrates under well-controlled operating variables. The thus-obtained anti-glare glasses were baked at 180°C for 1h. The properties of gloss, haze, transmittance, and surface roughness were measured and the results were summary in Table 1. The results indicate that the average transmittance of sol-gel films descends in the following order of methanol > propanol

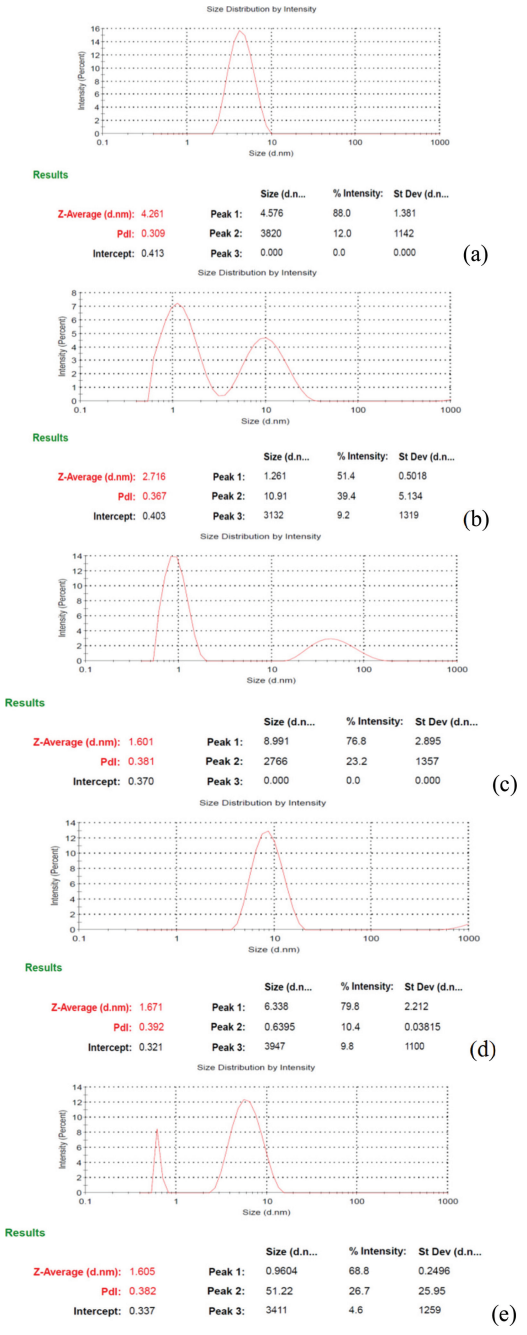

Figure 1. Charts of particle size analysis of silica sol-gel particles from different solvents of (a) methanol, (b) ethanol, (c) propanol, (d) isopropanol, and (e) n-butanol.

> ethanol > n-butanol > isopropanol, the average haze is methanol > ethanol > isopropanol > propanol > n-butanol, and the average gloss

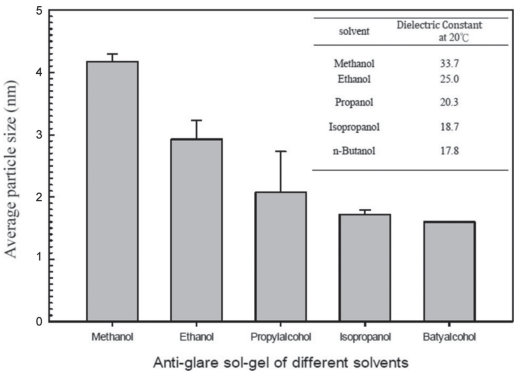

Figure 2. Average particle size of silica sol-gel with different alcohol solvents.

Table 1. Averaged gloss, haze, transmittance, and surface roughness of the prepared anti-glare silica sol-gel/glass samples with different alcohols.

Solvent	Average gloss (GU)	Average haze (%)	Average TT (%)	Average Ra (μm)
Blank*	165	0.09	90.30	0.013
Methanol	74.3	9.76	92.35	0.146
Ethanol	87.4	4.57	92.10	0.167
Propanol	104	3.56	92.12	0.104
Isopropanol	123.3	3.63	91.83	0.064
n-Butanol	100.4	2.92	91.92	0.172

*Blank glass without coating.

is isopropanol > propanol > n-butanol > ethanol > methanol. In summary, the solvent of methanol shows the highest average transmittance of 92.35%, maximum average haze value of 9.76%, and the lowest average gloss value of 74.3 GU among the applied alcohols. However, the solvent of isopropanol exhibits the minimum arithmetical mean deviation (Ra) of surface roughness (0.064 μm).

When compared with the surface anti-reflection images shown in Fig. 3, the silica sol-gel prepared from the solvent of methanol provides the most antiglare and transmittance performance as a result that we can clearly distinguish the back words. According to the literature (Liu, et al., 2002; Liu, et al., 2010; Liu, et al., 2011; Wang, et al., 2009a; Wang, et al., 2009b), poly (methyl methacrylate) beads, polystyrene beads, acrylic-styrene resin particles have been extensively employed as the anti-glare coating layer. Unfortunately, introducing these materials was usually associated with poor transmittance. However, in this study, increasing ant-glare arouses no significant depression in

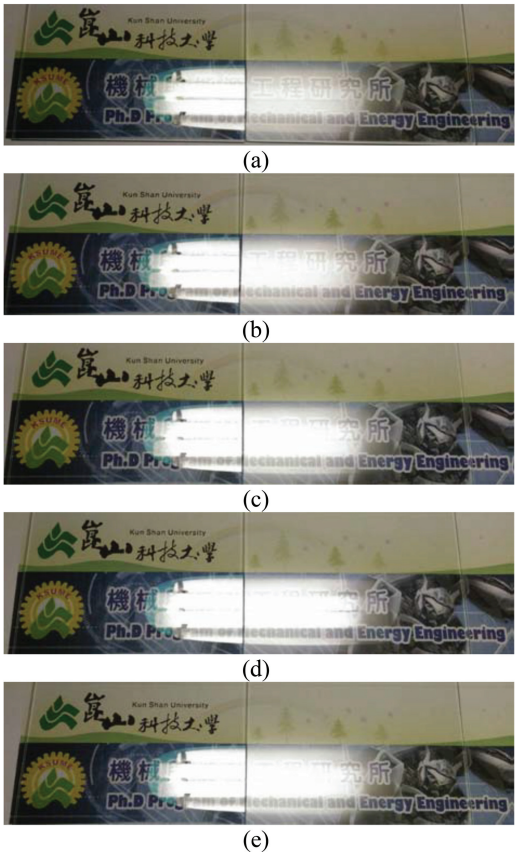

Figure 3. Reflection image of silica anti-glare glass samples prepared from (a) methanol, (b) ethanol, (c) propanol, (d) isopropanol, and (e) n-butanol by an auto spray-coating method.

transmittance, approximately maintaining at 92%, which can be ascribed to a lower reflective index of silica sol-gel thin film (~1.45) than that of glass substrates (~1.52) (Zhia, et al., 2018; Xi, et al., 2007; Kesmez, et al., 2011). In addition, surface roughness is independent of the observation of gloss and haze, and makes no noticeable influence on anti-glare performance.

4 CONCLUSIONS

In this paper, the synthesis of silica sol-gel nanoparticles and their application as anti-glare thin films on glass substrates with various alcohol solvents of methanol, ethanol, propanol, isopropanol, and nbutanol under acid-catalyzed aqueous systems have been investigated. The particle size of the synthesized silica sol-gel decreases in order of methanol > ethanol > propanol > isopropanol > n-butanol, which is likely to be caused by the steric effect during condensation reactions. Coating silica sol-gel particles on to glass substrates through the auto-spray systems can offer excellent anti-glare effects without hampering their transmittance. Among the alcohol solvents examined, methanol generates the best antiglare performance due to the lowest average gloss and the highest average haze. In addition, the surface roughness seems to exert insignificant impact on the anti-glare for silica sol-gel particles.

REFERENCES

Brinker, C.J., G.W. Scherer, 1990. *Sol-Gel Science: The Physics and Chemistry of Sol-Gel Processing*.

Chen, H., L. Chen, H. Gong, et al., 2016. *Anti-glare substrates with a uniform textured surface and low sparkle and methods of making the same*, WO2016069113A1.

Chou, K.S., C.C. Chen, 2003. *Advances in Technology of Materials and Materials Processing Journal*, 5 (2003) 31–35.

Kesmez, Ö., E. Burunkaya, N. Kiraz, et al., 2011. *Journal of Non-Crystalline Solids* 357 3130-3135.

Klein, L.C., 1994. *Sol-Gel Optics: Processing and Applications*.

Lin, P.L., Y.T. Lin, S.L. Hwang, et al., 2008. *Displays* 29 25–32.

Liu, B.T., Y.T. Teng, 2010. *Journal of Colloid and Interface Science* 350 421–426.

Liu, B.T., Y.T. Teng, R.H. Lee, et al., 2011. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 389 138–143.

Liu, W.C., C.C. Yang, W.C. Chen, et al., 2002. *Journal of Non-Crystalline Solid* 311 233.

Lugscheider, E., C. Barimani, P. Eckert, et al., 1996. *Computational Materials Science*, 7 109–114.

Ma, X.K., N.H. Lee, H.J. Oh, et al., 2010. *Colloids and Surfaces A: Physicochem. Eng. Aspects* 358 172–176.

Malay, O., I. Yilgor, Y.Z. Menceloglu, 2013. *Journal of Sol-Gel Science and Technology* 67 351–361.

Matsuyama, I., S. Satoh, M. Katsumoto, et al., 1991. *Journal of Non-Crystalline Solids* 135 22–28.

Wang, J.H., C.F. Chen, J.R. Ho, et al., 2009. *Optics & Laser Technology* 41 (2009) 804–808.

Wang, J.H., S.Y. Lien, J.R. Ho, et al., 2009. *Optical Materials* 32 374–377.

Xi, J.Q., M.F. Schbert, J.K. Kim, et al., 2007. *Nature Photonics* 1 176–179.

Zhia, J.H., L.Z. Zhanga, 2018. *Applied Surface Science* 454 239–248.

Communication science and engineering

Engineering Innovation and Design – Lam et al. (Eds)
© 2019 Taylor & Francis Group, London, ISBN 978-0-367-02959-3

Resource allocation in an LTE-A network for Machine Type Communications (MTC) over a Wi-Fi spectrum under an LAA framework

Rosabella Ika Yuanita & Ding-Bing Lin
Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan

Gamantyo Hendrantoro
Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

ABSTRACT: A Heterogeneous Network (HetNet) is a combination of networks, such as macro cell and small cell, which has been introduced as an LTE-Advanced (LTE-A) standard in 3GPP Release 10 in order to provide higher network capacity and coverage. However, interference could occur because macro cells have a much higher transmit power than small cells, and the cell-edge users of small cells may use the same radio resources as macro cell users. To cope with this problem, we propose an optimization of resource allocation using Taguchi's method to improve a system's throughput. In addition, Machine to Machine Communication (M2M), also known as Machine Type Communication (MTC), is a future wireless communication technology that will be used for industry 4.0. In order to provide MTC services, 3GPP allows some wireless networks, such as LTE-A, Wi-Fi and Internet-of-Things (IoT), to co-exist in the 5 GHz band through Carrier Aggregation (CA), namely Licensed Assisted Access (LAA). MTC devices will be allowed to access the channel with LTE-A users using the Listen Before Talk (LBT) protocol. Our simulation results show that Proportional Fair (PF) is the optimum scheduler that can be used in both licensed and unlicensed bands for the LTE-A network when Frame Based Equipment (FBE) based LBT is adopted as the channel access mechanism. Furthermore, the performance of Taguchi's method is better than a genetic algorithm because of its faster convergence speed.

1 INTRODUCTION

Machine to Machine (M2M) communications, also known as Machine Type Communications (MTC), will be the wireless communication technology used for industry 4.0 applications in the future. Unlike Human Type Communications (HTC), MTC does not require human intervention when transferring its data. By 2020, it is predicted that there will be about 50 billion IoT Devices (IoTD) connected to mobile devices worldwide (Starkloff, 2015). 3GPP provides MTC services in an LTE network. The rapidly increasing number of subscriber in LTE-A network affected on the traffic demand, it should be a key challenge for network operators to increase network capacity in order to fulfill the requirements of cellular users without degrading the Quality of Service (QoS). One of the solutions for network capacity enhancement, aside from adding a number of base stations, is Carrier Aggregation (CA). 3GPP has developed CA in the LTE-A network by combining the unlicensed spectrum and the licensed spectrum in the downlink using Listen Before Talk (LBT) protocol to provide a larger bandwidth system. This technology is called Licensed Assisted Access (LAA).

With regards to 3GPP Release 10, a Heterogeneous Networks (HetNet) environment is considered to be an LTE-A standard. HetNet has several serving cells, consisting of primary and secondary serving cells. When the component carriers are used in different frequency bands, a difference will be experienced in both coverage and pathloss, as seen in Figure 1. However, this might cause a problem, because macro cells have a much higher transmit power than small cells. The cell-edge User Equipments (UEs) in small cells can interfere with the neighboring macro cells and both cells may use the same radio resources in the different time ranges (subframes). To cope with this problem, we propose an optimization of resource allocation using Taguchi's method to allocate channels to users in order to achieve higher LTE throughput.

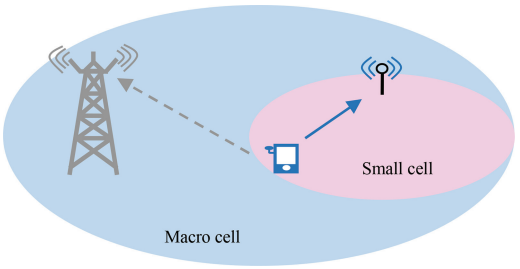

Figure 1. The cell-edge UE in a heterogeneous network has an inter-cell interference.

2 ADOPTED METHODS

2.1 *Taguchi's method*

Taguchi's method is an iterative optimization algorithm based on an Orthogonal Array (OA) concept. It has benefits for time efficiency, because although it does not require many experiments in order to solve problems that have many input parameters, it can still achieve a proper solution. The first procedure of the Taguchi method is the design of the OA construction, OA(*N, k, s, t*), where *N, k, s,* and *t* denote the number of experiments needed to be carried out per design iteration, the number of parameters that need to be optimized, corresponding values for the levels depending on the parameters, and various different applications and strengths. According to Iwamura et al. (2010), three levels are needed (s = 3) in order to characterize the non-linear effect and this is sufficient for each input with a strength of 2 (t = 2), which is efficient for most problems because the rows can result in a small number. Because of having three levels and the strength of two, the OA construction of OA (9, 4, 3, 2) should be chosen. The input parameters of this paper have four control factors, including the number of IoT devices (IoTD), bandwidth size, scheduler, and the length of the Channel Occupancy Time (COT). Each factor has three levels, as listed in Table 1. The amount of Physical Resource Block (PRB) in a bandwidth size is noted in Table 2. We then consider the system throughput performance as a fitness function, which can be expressed as (ETSI EN 301 893 V1.7.1):

$$TP_{1S} = N_{RB}.N_{RE\_PRB}.modBits.CR.TS_{1S} \quad (1)$$

where N_{RB} is the total number of Resource Blocks (RB) per time slot (0.5 ms), and N_{RE_PRB} is the number of Resource Elements (RE) per resource block. For a normal cyclic prefix, one RB contains 12 sub-carriers and 7 OFDM symbols, thus there are 84 resource elements in one RB. The *modBits* variable represents the number of bits carried by one symbol in a given modulation (i.e. for 64 QAM, modBits is equal to 6). Lastly, *CR* is the code rate of the error correction coding and TS_{1S} is a constant transforming the throughput from 0.5 ms scale to 1s scale. All of these values can be taken from Slanina et al. (2014) and the 3GPP Technical Report TR 36.213 (2010).

Table 1. Array table for optimization parameters.

Experiment	Number of IoTD	Bandwidth	Scheduler	COT
1	1	1	1	1
2	1	2	2	3
3	1	3	3	2
4	2	1	2	2
5	2	2	3	1
6	2	3	1	3
7	3	1	3	3
8	3	2	1	2
9	3	3	2	1

Table 2. Variation levels of optimization parameters.

Parameters \ Levels	1	2	3
Number of IoTD	20	50	100
Bandwidth	5	10	20
Scheduler	RR	BET	PF
COT	5 *ms*	7 *ms*	10 *ms*

2.2 *Genetic Algorithm (GA)*

A genetic algorithm is an heuristic search algorithm inspired by Darwin's theory of natural selection. The algorithm is solved iteratively, beginning with a randomly generated group composed of solutions represented by the number of chromosomes, which is also known as one population. By referring to the system's throughput comparison result when using Taguchi's method as an optimization algorithm, we can determine the Proportional Fair (PF) scheduler as a fitness function in the evaluation phase. We can then perform the selection process by using roulette wheel selection to produce offsprings. This will pick the chromosome with high probability, which is proportionate to its fitness. If the random number generation is smaller than the crossover probability, then the next step will be to undergo the crossover procedure. The last step in the GA is mutation, which occurs when the random number generation is smaller than the mutation probability. The algorithm can be stopped if it no longer produces a greater improvement in the fitness of the best

Table 3. GA simulation parameters.

Parameters	Characterization/Values
Population size	200
Number of genes	50
Crossover method	One-point crossover
Crossover rate	0.9
Selection method	Roulettewheel
Mutation method	Order changing mutation
Mutation rate	0.001
Termination criterion	Improvement of fitness value

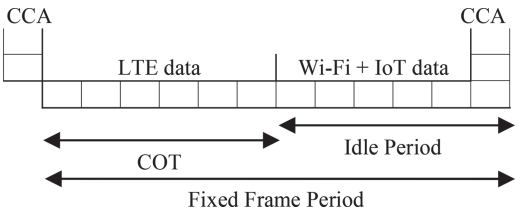

Figure 2. Adopted LTE-LAA frame structure and timing.

individual for successive iterations. Table 3 represents the GA simulation parameters.

2.3 *FBE-based Listen Before Talk (LBT)*

In the unlicensed band, we cannot adopt an LTE-LAA based Orthogonal Frequency Division Multiple Access (OFDMA) frame structure because both the Wi-Fi and IoT networks are also utilized. Frame Based Equipment (FBE) based LBT is used according to a fixed frame structure, which is divided into a COT and idle period (Cox, 2012). LTE-LAA occupies the channel during the COT and remains silent, while the other networks occupy it during the idle period. All of the equipment should perform a Clear Channel Assessment (CCA) only once in each Fixed Frame Period (FFP) before accessing the channel for a channel observation time ≥ 20 μs, using Energy Detection (ED) to check the channel status. If the channel is clear, the equipment may transmit immediately for a COT of between [1, 10] ms. Otherwise, it should wait until it senses that the channel is idle and then transmit at the next FFP. A COT can be set for up to 10 subframes for LTE data transmissions. The minimum length of the idle period should be at least 5% of the COT. The adopted LTE-LAA frame structure and timing in our paper is illustrated in Figure 2. We assumed that the length of FFP is 20 ms. Hence, the duration of the idle period is 15 ms. The LAA bandwidth is fixed at 20 MHz.

3 SCHEDULING SCHEMES

3.1 *Blind Equal Throughput (BET)*

Blind equal throughput provides equal throughput to all UEs and neglect their Channel Quality Index (CQI). Therefore, this scheme does not require feedback from the UE, and the Modulation and Coding Scheme (MCS) will be given to the user according to the wideband CQI. However, BET utilizes the past average throughput of the user in order to schedule the user. The priority metric for BET is given by:

$$m_{i,k}^{BET} = \arg\max_i \frac{1}{\overline{R^i}(t-1)} \tag{2}$$

where $\overline{R^i}(t-1)$ is the past average throughput of user i at time t using an exponential moving average. This scheme recalculates and updates the expected average throughput of the UE at every Transmission Time Interval (TTI) and compares this with the past average throughput of other UEs. If a user is no longer the lowest, the Scheduling Blocks (SB) will be assigned to another user who has the highest priority until a better average throughput is achieved.

3.2 *Round Robin (RR)*

Round robin is the simplest and fairest algorithm for resource allocation since users will be scheduled for equal amounts without any priority and radio channel quality is not considered. MCS will be allocated to users based on wideband CQI. Hence, a user is not required to send feedback to eNodeB. If all users have been scheduled in the first TTI, they will be placed to the back of the queue for the next TTI and be scheduled in circular order.

3.3 *Proportional Fair (PF)*

Proportional fair is the most common scheduler used in a cellular network because of the trade-off between fairness and the throughput performance index. Unlike the BET and RR schedulers, this scheme takes CQI into account. Users who have good channel conditions have a higher priority to be scheduled. MCS is given to users based on their UE feedback to eNodeB with 4-bit CQI, as listed in Slanina et al. (2014). In principle, a user can be assigned to various scheduling blocks (SBs), which can be either adjacent or non-adjacent depending on the present channel conditions and past throughput performance Ri(t) at the end of subframe t, calculated by using an exponential moving average. The priority metric of a PF scheduler can be defined as:

$$m_{i,k}^{PF} = \arg\max_i \frac{d_k^i}{\overline{R^i}(i-1)} \quad (3)$$

where d_k^i is the achievable data rate of user i on the k-th PRB.

4 PERFORMANCE EVALUATION

4.1 *Network model*

MATLAB is the simulation platform used to evaluate the LTE-A system throughput performance of three schedulers, including round robin, blind equal throughput and proportional fair, in the 20 MHz of bandwidth size. The simulation lengths of the licensed and unlicensed LTE-A are 50 TTIs and 10 FFPs, respectively. There are 50 users in the licensed band and 10 users in the unlicensed band. The simulation parameters are summarized in Table 4 and the user distribution in the heterogeneous network layout is illustrated in Figure 3.

In our simulation network, co-existing networks, including macro cell Base Station (BS), small cell BS and Wi-Fi, are located in the center of the hexagonal grid. Wi-Fi is used to provide a service for IoT devices. Small cell users and IoT devices are distributed in the center of the hexagon near to the access point and small cell BS, meanwhile the macro cell users are distributed at the edge of the hexagon. Macro cells are the Primary Component Carrier (PCC) and small cells are the Secondary Component Carrier (SCC). Edge cell and small cell users use 2 GHz and 5.2 GHz of frequency, respectively, when they want to communicate with each other.

Table 4. Simulation parameters.

Parameters	Characterization/Values
User arrival dist.	Random and uniform distribution
Antenna gain	Macro cell: 18 dB Small cell: 2 dB
Tx power	PSC: 48 dBm SSC: 18 dBm
Channel model	**Licensed-Urban** • Macro cell UE to macro cell BS: ITU UMa • Macro cell UE to small cell BS: Winner II **Licensed-Suburban** • Macro cell UE to macro cell BS: Cost-231 Hata • Macro cell UE to small cell BS: Winner II **Unlicensed**: Industrial Indoor Channel (Tanghe et al., 2008)
Wi-Fi standard	802.11n
Number of IoTD	20

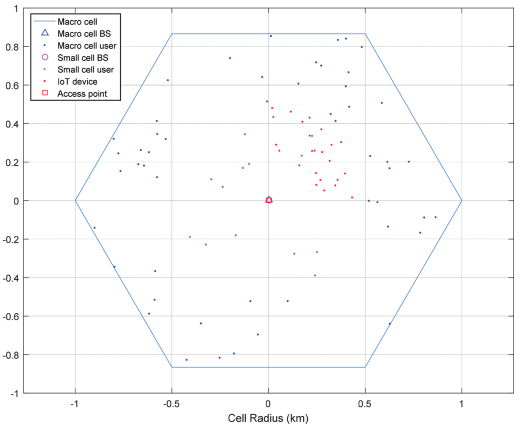

Figure 3. User distribution in the heterogeneous network layout.

4.2 *Results and discussions*

Figure 4 shows a system throughput comparison using different schedulers. Among the three kinds of scheduler shown, PF achieves the highest throughput in both the licensed and unlicensed bands. The reason for this is because PF takes CQI into account, while RR and BET do not. However, the PF scheduler has benefits for users who are close to the eNodeB, because their Signal-to-Interference-plus-Noise Ratio (SINR) is higher than that of edge users who may have a lower probability of being scheduled, which therefore leads to starving and low throughput for edge users. In addition, when comparing between licensed and unlicensed bands, the system throughput for unlicensed bands is higher than that for licensed bands because of the feature of carrier aggregation in unlicensed bands. Also, the user position in unlicensed bands is closer to the eNodeB.

We also evaluated the performance between Taguchi's method and genetic algorithms (GA), as illustrated in Figure 5. The result shows that the convergence in GA is achieved at the ninth iteration. Therefore, Taguchi's method performs better than GA because Taguchi uses OA to find the global optima quickly and does not utilize any steps, unlike GA. It can therefore be used to analyze many different parameters without requiring a large number of experiments. Meanwhile, GA uses random initialization, and thus there is no guarantee of achieving a global optima. Furthermore, the fitness function, crossover and mutation rate of GA need to be determined properly, because

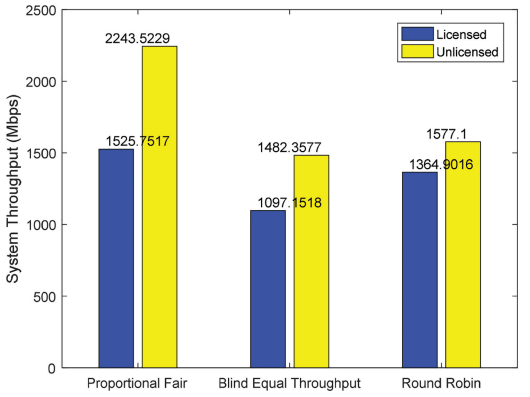

Figure 4. System throughput comparison of different schedulers.

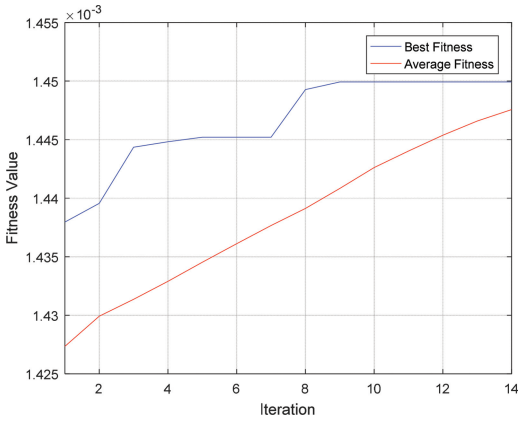

Figure 5. Fitness graph of GA in every iteration.

Table 5. Convergence speed comparison between Taguchi and GA.

	Optimization algorithm	
	Taguchi's method	Genetic algorithm
Number of iterations	1	9
Running time	0.356 s	1.82 s

an infeasible value can decrease the performance so that eventually the optimum value could not be met. As listed in Table 5, the convergence speed of GA is slower, and the running time is also longer than that of Taguchi.

5 CONCLUSION

After we observed and analyzed the system throughput, PF was found to be the optimum scheduler as it can be used in both unlicensed and licensed bands for small cell and macro cell users. FBE-based LBT can be used for UEs as an access method in the unlicensed bands, but this may be a problem when all of the networks are in co-existence because they may have to wait for a long time to occupy the channel. Also, compared to GA, Taguchi's method is an appropriate algorithm for resource scheduling cases in both licensed and unlicensed bands.

REFERENCES

3GPP Technical Report TR 36.213. (2010). *Physical Layer Procedures* v 9.2.0.

Cox, C. (2012). *An introduction to LTE: LTE, LTE-Advanced, SAE and 4G mobile communications. Wiley*. UK.

Iwamura, M., Etemad, K., Fong, M.H., Nory, R., and Love, R. (2010). *Carrier Aggregation Framework in 3GPP LTE-Advanced. IEEE Commun. Mag.,* 48, 60–67.

ETSI EN 301 893 V1.7.1.

Slanina, M., Klozar, L. & Hanus, S. (2014). *Practical Measurement of Data Throughput in LTE Network Depending on Physical Layer Parameters. The 24th IEEE International Conference Radioelektronika.*

Starkloff, E, (2015). *Business 2 Community*. http://www.business2community.com/tech-gadgets/future-5ginternet-everyone-everything-01303812#ObtcAoJEKdfCUqZs.97.

Tanghe, E., Joseph, W. Verloock, L., Martens, L., Capoen, H., Herwegen, K.V., and Vantomme, W. (2008). *The industrial indoor channel: Large scale and temporal fading at 900, 2400, and 5200 MHz. IEEE Transactions on Wireless Communications.* 7. 2740–2751.

Engineering Innovation and Design – Lam et al. (Eds)
© 2019 Taylor & Francis Group, London, ISBN 978-0-367-02959-3

An efficient hybrid direction of arrival estimation scheme for massive multiple-input multiple-output systems

Ann-Chen Chang
Department of Information Technology, Ling Tung University, Taichung, Taiwan

Wei Jhang
PhD Program of Electrical and Communications Engineering, Feng Chia University, Taichung, Taiwan

Shiaw-Wu Chen
Department of Automatic Control Engineering, Feng Chia University, Taichung, Taiwan

ABSTRACT: In this paper, an efficient hybrid Direction of Arrival (DOA) estimation scheme based on the orthogonal projection criterion is devised for a massive uniform linear array. In this scheme, the DOA estimator, based on a Discrete Fourier Transform (DFT), is first applied to acquire coarse initial DOA estimates for a single data snapshot. Then, the fine DOA is accurately estimated through using the iterative search estimator within a very small region. It iteratively searches for the correct DOA vector by maximizing the objective function using a Taylor series approximation of the DOA vector with the one initially estimated. Since the proposed scheme does not need to perform eigendecomposition and a spectrum search while maintaining better DOA estimates, it also has low complexity and real-time capability. Simulation results are presented to demonstrate the efficiency of the proposed scheme.

1 INTRODUCTION

A Multiple-Input Multiple-Output (MIMO) system utilizes multiple antennas to transmit signals and detect objects. It has widespread application in the radar and wireless communication fields. As a promising technology in 5G, massive MIMO, which is also known as the large-scale MIMO, has been attracting significant attention owing to its unprecedented potential for high-spectral efficiency (Shu et al., 2018). It is well known that the beam forming/precoding performance relies closely on the accurate channel state information estimation, or more specifically, the accurate Direction Of Arrival (DOA) estimation (Yao et al., 2017). So this requires an accurate DOA estimation method, which becomes a more crucial task in massive MIMO systems.

The DOA estimation has been of interest to the signal processing community for decades. Enormous interest has been raised within the array processing community in the high-resolution subspace-based angle estimation algorithms, which include Multiple Signal Classification (MUSIC) (Yang et al., 2014), Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) (Hu et al., 2014), and their variants (Yao et al. 2017). However, these conventional MUSIC and ESPRIT algorithms are not suitable for the massive MIMO systems. They both need Eigenvalue Decomposition (EVD) of the autocorrelation matrix and/or spectrum peak search, which have high computational complexity. The subspace-based methods require the formulating of a reliable signal-subspace or noise-subspace method. However, a huge amount of computation would be involved, where a large array element size is required in applications. A possible alternative to the subspace-based method for DOA estimation with computational saving is the Orthogonal Projection (OP) method (Chang & Shen, 2013). Meanwhile, the computational requirements are very high for the searching-based estimators, where the complexity and estimation accuracy strictly depend on the grid size used during the search. It is time-consuming and the search grid is ambiguous.

To lessen the above-mentioned problem, this paper presents a hybrid DOA estimation scheme, based on OP criterion without EVD for massive Uniform Linear Array (ULA) systems. Specifically, the proposed hybrid scheme first obtains initial DOA estimates through Discrete Fourier Transform (DFT) and then finds accurate fine DOA estimates via the iterative search approach within each small region. The basic idea of the fine estimator is based on the principle that the output power of an

estimator achieves a local maximum if the direction vector coincides with the desired signal. It iteratively searches for a correct DOA vector, by maximizing the objective function with OP criterion using a first-order Taylor series approximation of the DOA vector with the one initial estimate given. Let the number of signals be S. After obtaining initial DOA estimates, the whole possible DOA searching range forms S smaller search ranges. The maximum value of the objective function is searched in each smaller range to get each signal's DOA, rather than finding the S peaks in the whole range, respectively. Therefore, the iterative fine estimator can conquer the effect of ambiguous peaks with initial DOA estimates, and reduce the complexity greatly with a self-tuned search grid. It is more suitable for fine DOA searching in the actual operations. Simulation results are provided to demonstrate the effectiveness of the proposed hybrid scheme.

2 PRELIMINARIES

2.1 *Signal model*

Consider a massive ULA with M isotropic sensors. There are $S(S<M)$ uncorrelated narrowband source signals impinging on the array from directions $\{\theta_1,\theta_2,\cdots,\theta_S\}$. The $M\times 1$ array observation vector is given by:

$$\mathbf{y}(t)=\mathbf{A}(\theta)\boldsymbol{\beta}(t)+\mathbf{n}(t),\ t=1,2,\cdots,N \tag{1}$$

where $\mathbf{A}(\theta)=[\mathbf{a}(\theta_1),\ \mathbf{a}(\theta_2),\ \cdots,\mathbf{a}(\theta_S)]$ is the steering matrix with size $M\times S$ and the steering vector due to the sth source can be expressed as $\mathbf{a}(\theta_s)=[1,\ e^{j2\pi d\sin(\theta_s)/\lambda},\ \cdots,\ e^{j2\pi d(M-1)\sin(\theta_s)/\lambda}]^T$ with $(\bullet)^T$ denoting the transpose. For ULA, the interelement spacing is $d=0.5\lambda$ and λ is the carrier wavelength. $\boldsymbol{\beta}(t)=[\beta_1(t),\beta_2(t),\cdots,\beta_S(t)]^T$ is the source signal vector and N is the number of available data snapshots. $\mathbf{n}(t)$ is an $M\times 1$ additive white Gaussian noise vector with zero mean and covariance $\sigma_n^2\mathbf{I}_M$, where $\mathbf{I}_M$ is the $M\times M$ identity matrix. Assume that the noise is uncorrelated with the signal $\mathbf{s}(t)$. The $M\times M$ autocorrelation matrix of the received observation vector $\mathbf{y}(t)$ is given by:

$$\begin{aligned}\mathbf{R}&=E\{\mathbf{y}(t)\mathbf{y}^H\}\\&=\mathbf{A}(\theta)\mathbf{R}_\beta\mathbf{A}^H(\theta)+\sigma_n^2\mathbf{I}_M\end{aligned} \tag{2}$$

where $\mathbf{R}_\beta=E\{\boldsymbol{\beta}(t)\boldsymbol{\beta}^H(t)\}$, $E\{\bullet\}$ denotes expectation operator, and $(\bullet)^H$ represents the conjugate transpose.

2.2 *Searching-based OP estimator*

First, the steering matrix $\mathbf{A}(\theta)$ with appropriate partition is defined as $\mathbf{A}(\theta)=[\mathbf{A}_1^T(\theta)\ \ \mathbf{A}_2^T(\theta)]^T$, where $\mathbf{A}_1(\theta)$ and $\mathbf{A}_2(\theta)$ are the $S\times S$ and $(M-S)\times S$ submatrices, respectively. The relationship between $\mathbf{A}_1(\theta)$ and $\mathbf{A}_2(\theta)$ is $\mathbf{A}_2(\theta)=\mathbf{U}^H\mathbf{A}_1(\theta)$, where the transformation matrix $\mathbf{U}=[\mathbf{A}_2(\theta)\mathbf{A}_1^{-1}(\theta)]^H$. Define $\Phi=[\mathbf{U}^H-\mathbf{I}_{M-S}]^H$, then $\Phi^H\mathbf{A}(\theta)=\mathbf{0}$. Here, we use the received signal to estimate $\boldsymbol{\Phi}$. In the absence of noise, the $M\times M$ matrix $\mathbf{R}$ with appropriate partition is defined as $\mathbf{R}=[\mathbf{R}_1\,|\,\mathbf{R}_2]$, where $\mathbf{R}_1$ consists of the first S columns of $\mathbf{R}$ and $\mathbf{R}_2$ consists of the last $M-S$ columns of $\mathbf{R}$. It has a linear relationship between $\mathbf{R}_1$ and $\mathbf{R}_2$, and $\mathbf{R}_2=\mathbf{R}_1\mathbf{U}$. The matrix $\mathbf{U}$ can be estimated using the least squares method. This is achieved by minimizing the $\|\mathbf{R}_2-\mathbf{R}_1\mathbf{U}\|^2$, and $\|\bullet\|$ indicates the Euclidean norm. Then, the estimated $\hat{\mathbf{U}}=\mathbf{R}_1^+\mathbf{R}_2$, where $(\bullet)^+$ is the pseudo inverse operation and $\mathbf{R}_1^+=(\mathbf{R}_1^H\mathbf{R}_1)^{-1}\mathbf{R}_1^H$. Therefore, the estimated noise subspace matrix is given by:

$$\hat{\Phi}=[\hat{\mathbf{U}}^H-\mathbf{I}_{M-S}]^H \tag{3}$$

Due to $\boldsymbol{\Phi}$ being orthogonal to $\mathbf{A}(\theta)$ such that $\Phi^H\mathbf{A}(\theta)=\mathbf{0}$, the columns in $\hat{\boldsymbol{\Phi}}$ span the same noise subspace and exploit this to develop a MUSIC-like DOA estimator without using eigendecomposition. However, the column vectors of $\hat{\boldsymbol{\Phi}}$ cannot form an orthonormal basis; we use orthogonal projection to this subspace to satisfy the orthonormality and improve the performance through the base of the orthonormality and zero-kernel space. Therefore, the projection matrix based on OP is $\mathbf{P}_\Phi=\hat{\boldsymbol{\Phi}}(\hat{\boldsymbol{\Phi}}^H\hat{\boldsymbol{\Phi}})^{-1}\hat{\boldsymbol{\Phi}}^H$. Then, θ_s corresponds to the largest local maximum of $J(\theta)$ with the searching grid μ_1 and is given by:

$$J(\theta)=[\mathbf{a}^H(\theta)\mathbf{P}_\Phi\mathbf{a}(\theta)]^{-1} \tag{4}$$

where $\mathbf{a}(\theta)=[1,\ e^{j2\pi d\sin(\theta)/\lambda},\cdots,\ e^{j2\pi d(M-1)\sin(\theta)/\lambda}]^T$ is the direction vectors associated with the actual signals and $\theta\in[-90°,90°]$ varies within the whole searching space. Here, the searching-based approach has to scan the complete degree interval.

3 PROPOSED HYBRID DOA ESTIMATION SCHEME

3.1 *DFT-based initial DOA estimator*

DFT is one of the conventional non-parametric spectrum analysis methods that have much lower resolution than their parametric counterparts (Cao et al., 2017). Define the normalized $M\times M$ DFT matrix $\mathbf{F}$ with its (p,q)th element given by $[\mathbf{F}]_{p,q}=\frac{1}{\sqrt{M}}e^{-j2\pi pq/M}$. Meanwhile, the DFT of a steering vector $\mathbf{a}(\theta_s)$ is $\tilde{\mathbf{a}}(\theta_s)=\mathbf{F}\mathbf{a}(\theta_s)$ and the qth element of $\tilde{\mathbf{a}}(\theta_s)$ is:

$$[\tilde{\mathbf{a}}(\theta_s)]_q = \frac{1}{\sqrt{M}} \frac{\sin\left[\frac{M}{2}\left(\frac{2\pi q}{M} - \frac{2\pi d \sin\theta_s}{\lambda}\right)\right]}{\sin\left[\frac{1}{2}\left(\frac{2\pi q}{M} - \frac{2\pi d \sin\theta_s}{\lambda}\right)\right]} e^{-j\frac{M-1}{2}\left(\frac{2\pi q}{M} - \frac{2\pi d \sin\theta_s}{\lambda}\right)} \tag{5}$$

If the system has an infinite number of antennas (i.e. M approaches to infinity), there always exists an integer $q = \frac{M}{2}\sin(\theta_s)$, such that $[\tilde{\mathbf{a}}(\theta_s)]_q = \sqrt{M}$, while the other elements of $\tilde{\mathbf{a}}(\theta_s)$ are all zero. It means that $\tilde{\mathbf{a}}(\theta_s)$ is ideally sparse and all powers are concentrated on the qth DFT point. However, a practical array aperture cannot be infinitely large, even if M could be as great as hundreds or thousands in a massive ULA. In this case, $\frac{M}{2}\sin(\theta_s)$ will not be an integer for the most time, and the power will 'leak' from the $\lfloor \frac{M}{2}\sin(\theta_s) \rceil$th DFT point to the nearby points, where $\lfloor x \rceil$ is a round operator that returns the nearest integer to x. Clearly, the degree of leakage is inversely proportional to M but is proportional to the deviation $(\frac{M}{2}\sin(\theta_s) - \lfloor \frac{M}{2}\sin(\theta_s) \rceil)$. Since $M >> 1$ for a massive antenna array, $\tilde{\mathbf{a}}(\theta_s)$ can still be approximated as a sparse vector with most of the power concentrated around the $\lfloor \frac{M}{2}\sin(\theta_s) \rceil$ element. Hence, the peak power position of $\tilde{\mathbf{a}}(\theta_s)$ is still useful to indicate an initial DOA estimate.

Then, the DFT of $\mathbf{y}(t)$ can be formulated as:

$$\mathbf{x} = \mathbf{F}\mathbf{y}(t) \tag{6}$$

with $[\mathbf{x}]_q = \sum_{s=1}^{S} [\tilde{\mathbf{a}}(\theta_s)]_q \beta_s(t) + [\mathbf{F}\mathbf{n}(t)]_q$. So we can locate the S largest peaks of magnitude spectrum $|\mathbf{x}|$, denoted as $\{\hat{q}_s\}_{s=1}^{S}$, and obtain the coarse initial DOA estimates as:

$$\hat{\theta}_s^0 = \sin^{-1}(2\hat{q}_s/M),\ s = 1, 2, \ldots, S \tag{7}$$

It is noted that the resolution of $\sin(\hat{\theta}_s^0)$, by directly applying DFT, is still limited by half of the DFT interval (i.e. $1/(2M)$). After the initial DOA $\hat{\theta}_s^0$ is obtained, the searching region from $[-\frac{\pi}{2}, \frac{\pi}{2}]$ reduces to $[-\sin^{-1}(\frac{1}{M}), \sin^{-1}(\frac{1}{M})]$. Nevertheless, it is noted that the proposed iterative fine estimator only needs to search within a much smaller region.

3.2 *Iterative fine DOA estimator*

As we know, the searching-based OP estimator (Chang & Shen, 2013) needs exhaustive search with fairly high computational complexity in order to implement the high-resolution DOA estimation. A similar method had been used in Er and Ng (1994) for pointing error correction in adaptive beam forming. According to Equation 4, the objective function can be rewritten as:

$$f(\theta) = \underset{\theta}{\text{Min}} \left| \mathbf{a}^H(\theta)\mathbf{P}_\Phi\mathbf{a}(\theta) \right| \tag{8}$$

To locate the optimal shifted phase, one can simply search $\Delta\theta_s^i$ over the very small region $[-\sin^{-1}(\frac{1}{M}), \sin^{-1}(\frac{1}{M})]$ at the ith iteration. For the DOA estimation of the sth source signal, the search grid $\Delta\theta_s^i$ is not fixed, and it is updated in each iteration processing. At the ith iteration, let the first deviation of $\mathbf{a}(\theta_s^i)$ be $\mathbf{a}^{(1)}(\theta_s^i) = \partial\mathbf{a}(\theta)/\partial\theta|_{\theta=\theta_s^i}$. A first-order Taylor series approximation of the steering vector $\mathbf{a}(\theta_s^i)$ is given by:

$$\mathbf{a}(\theta_s^{i+1}) = \mathbf{a}(\theta_s^i + \Delta\theta_s^i) \approx \mathbf{a}(\theta_s^i) + \Delta\theta_s^i\mathbf{a}^{(1)}(\theta_s^i) \tag{9}$$

where θ_s^{i+1} is the estimated effective DOA at the $(i+1)$th iteration. According to Equation 9, Equation 8 can be rewritten as:

$$f(\Delta\theta_s^i) = \underset{\Delta\theta_s^i}{\text{Min}} \left| [\mathbf{a}(\theta_s^i) + \Delta\theta_s^i\mathbf{a}^{(1)}(\theta_s^i)]^H\mathbf{P}_\Phi[\mathbf{a}(\theta_s^i) + \Delta\theta_s^i\mathbf{a}^{(1)}(\theta_s^i)] \right| \tag{10}$$

The solution of Equation 10 is given by:

$$\Delta\theta_s^i = -\frac{\text{Re}\{\mathbf{a}^H(\theta_s^i)\mathbf{P}_\Phi\mathbf{a}^{(1)}(\theta_s^i)\}}{[\mathbf{a}^{(1)}(\theta_s^i)]^H\mathbf{P}_\Phi\mathbf{a}^{(1)}(\theta_s^i)} \tag{11}$$

where $\text{Re}\{\bullet\}$ represents the real part of the complex quantity. We can update θ_s^i for the next iteration as follows:

$$\theta_s^{i+1} = \theta_s^i + \Delta\theta_s^i \tag{12}$$

At the beginning of the iteration processing, the initial DOA estimate $\hat{\theta}_s^0$ for each source signal can be set as Equation 7. Let μ_2 be a terminating error value and it can be set to a desired precision (similar to the search resolution in an interested range). At the ith iteration, if the estimating deviation $|\Delta\theta_s^i| \le \mu_2$, we think $\Delta\theta_s^i$ is small enough, the iteration process ends, and θ_s^i is the expected value. Note that the number of iterations will be different for each signal due to different a DOA condition. In fact, this processing is equal to search θ, which can make the derivative of the objective function (Equation 9) be zero. In each iteration, $\Delta\theta_s^i$ is computed first. If $\Delta\theta_s^i$ is very small, we consider the variance rate of the objective function to achieve stabilization, so it is unnecessary to increase the iteration number, and θ_s^i is taken as the desired value, otherwise the iterative processing will continue.

4 SIMULATION RESULTS

This section reports simulation results, demonstrating the effectiveness of the proposed scheme for DOA estimation. Consider a 128-element ULA with half-wavelength spacing. The symbols

used for DOA estimation are modulated with binary phase shift keying modulation, and average received signal power from all sources is the same. There are 12 equal power signals impinging on the array from: {−55°, −45°, −35°, −25°, −15°, −5°, 10°, 20°, 30°, 40°, 50°, 60°}. The Signal-to-Noise Ratio (SNR) is defined as the ratio of the power of the source signal to that of the additive noise. The Root Mean Square Error (RMSE) is defined as $\mathrm{RMSE} = \frac{1}{\Pi}\frac{1}{S}\Sigma_{\rho=1}^{\Pi}\sqrt{\Sigma_{s=1}^{S}\left(\theta_s - \hat{\theta}_s^{\rho}\right)^2}$, where Π indicates the independent simulation runs. Each of the simulation results presented is after N = 25 snapshots have been processed and is the average of $\Pi = 10^3$ independent runs with an independent noise sample for each run.

In the first example, the general behaviors of the DFT-based coarse estimator and iterative fine estimator are evaluated. Figure 1 shows an ideal magnitude spectrum of the DFT-based estimator without leakage under $\mathrm{SNR} = 10\mathrm{dB}$ and $\theta = \{-23°, 41°\}$. It is noted that the $\{q_s\}_{s=1}^{S}$ can be immediately estimated from the S largest peaks of magnitude spectrum $|\mathbf{x}|$. On the contrary, Figure 2 depicts the leakage existing magnitude spectrum of the DFT-based estimator under $\mathrm{SNR} = 10$ dB and $\theta = \{-23.3°, 41.5°\}$. It is shown in Figure 2 that the power will 'leak' from the $\left\lfloor \frac{M}{2}\sin(\theta_s) \right\rceil$th DFT point to the nearby points. However, the peak power position of magnitude spectrum $|\mathbf{x}|$ is still useful to indicate an index of the initial DOA estimate. Figure 3 depicts the RMSE versus the terminating error μ_2 for the iterative fine estimator in a different SNR. It demonstrates that different

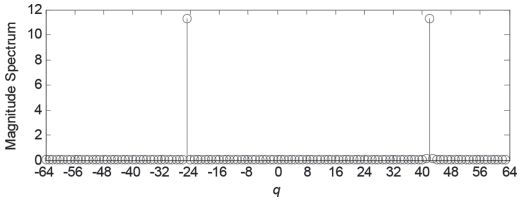

Figure 1. Ideal DFT spectrum of DFT-based estimator without leakage.

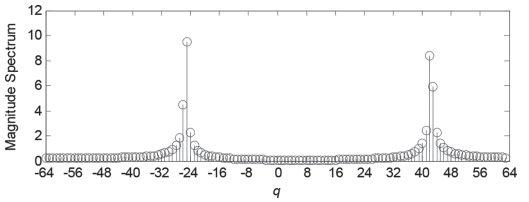

Figure 2. DFT spectrum of DFT-based estimator with leakage existing.

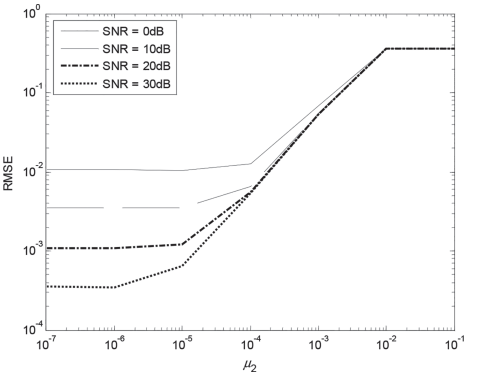

Figure 3. RMSE versus terminating error for the iterative fine estimator.

terminating error has better estimation resolution. A proper choice is $\mu_2 = 10^{-6}$ in the scenario range [0 dB, 30 dB].

In the second example, the performances of the proposed hybrid scheme is compared with the DFT-based (Cao et al., 2017), ESPRIT (Hu et al., 2014), MUSIC (Yang et al., 2014), OP (Chang & Shen, 2013), MVDR (Vaidyanathan & Buckley, 1995), and root-version estimators (Van Trees, 2002). The searching grid μ_1 is equal to 0.001°. Figure 4 presents RMSE versus number of snapshots under SNR = 10 dB. Clearly, increasing the number of snapshots induces performance improvement for all estimation methods, except the DFT-based estimator. Figure 5 shows RMSE versus SNR. We noted that the performance of the proposed hybrid scheme is very close to that of MUSIC, OP, root-MUSIC, and root-OP. Meanwhile, it has a lower complexity than does MVDR, MUSIC, and OP.

However, the proposed advantage of the iterative fine estimator is the reduction of computational complexity by reducing the number of searches while maintaining comparable performance. In practice, let us now compare the computational complexities of the OP and proposed hybrid scheme. Briefly, the searching efficiency is listed in Table 1. First, let the searching grid of OP be $\mu_1 = 0.001°$; then the number of searches $F_1 = 180{,}001$. Second, let the terminating error value of the iterative fine estimator be $\mu_2 = 10^{-6}$. Each signal has its own DOA, and the larger the DOA estimating deviation is, the higher is the search number required. The averaged numbers of searching were defined as $F_2 = \frac{1}{\Pi}\Sigma_{\rho=1}^{\Pi}\Sigma_{s=1}^{S}F_{2,s}^{\rho}$ with S = 12. It is noted that the proposed hybrid scheme is fast adaptively corrected from the previous iteration stage. Again, it is clear that the proposed hybrid scheme has higher searching efficiency.

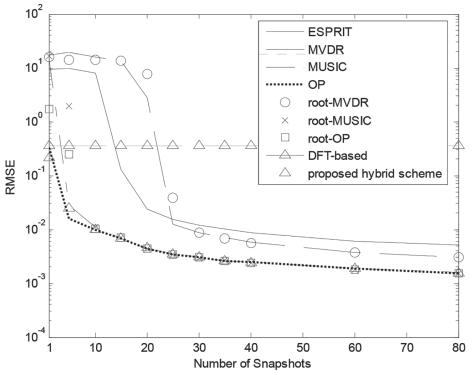

Figure 4. RMSE versus number of snapshots.

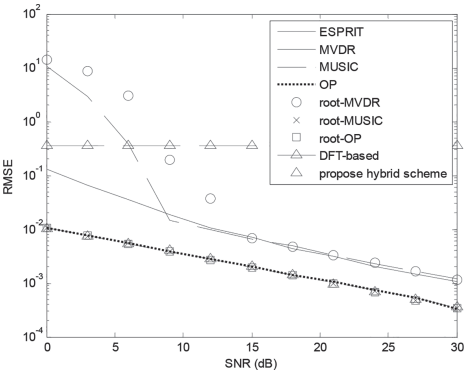

Figure 5. RMSE versus SNR.

Table 1. Searching efficiency of the iterative fine searching estimator.

SNR (dB)	0	3	6	9	12	15	18	21	24	27	30
F_2	6,936	6,290	5,948	5,773	5,664	5,604	5,562	5,535	5,511	5,490	5,471

5 CONCLUSIONS

This paper has presented a low complexity DOA estimation method for massive ULA. It is noted that the proposed hybrid estimation scheme does not perform EVD to obtain the noise subspaces. It is first found that the low resolution DFT spectrum becomes effective in providing good initial DOA estimation information. The iterative approach is used to solve search grids for the conventional spectral searching estimator. Several numerical results show that it has lower complexity than traditional methods, and the performance is better as compared to others.

ACKNOWLEDGMENT

This research was supported by the Ministry of Science and Technology under grant number MOST 107-2221-E-275-005.

REFERENCES

Cao, R., Liu, B., Gao, F. & Zhang, X. (2017). A low-complex one-snapshot DOA estimation with massive ULA. *IEEE Communications Letters*, *21*(5), 1071–1074.

Chang, A.C. & Shen, C.C. (2013). Orthogonal projection DOA estimation with a single snapshot. *IEICE Transactions on Communications*, *96*(5), 1215–1217.

Er, M.H. & Ng, B.C. (1994). A new approach to robust beamforming in the presence of steering vector errors. *IEEE Transactions on Signal Processing*, *42*(7), 1826–1829.

Hu, A., Lv, T., Gao, H., Zhang, Z. & Yang, S. (2014). An ESPRIT-based approach for 2-D localization of incoherently distributed sources in massive MIMO systems. *IEEE Journal of Selected Topics in Signal Processing*, *8*(5), 996–1011.

Shu, F., Qin, Y., Liu, T., Gui, L., Zhang, Y., Li, J. & Han, Z. (2018). Low-complexity and high-resolution DOA estimation for hybrid analog and digital massive MIMO receive array. *IEEE Transactions on Communications*, *66*(6), 2487–2501.

Vaidyanathan, C. & Buckley, K.M. (1995). Performance analysis of the MVDR spatial spectrum estimator. *IEEE Transactions on Signal Processing*, *43*(6), 1427–1437.

Van Trees, H.L. (2002). *Optimum Array Processing*. New York: John Wiley.

Yang, K.Y., Wu, J.Y. & Li, W.H. (2014). A low-complexity direction-of-arrival estimation algorithm for full-dimension massive MIMO systems. *IEEE International Conference on Communication Systems* (pp. 472–476).

Yao, B., Zhang, W. & Wu, Q. (2017). Weighted subspace fitting for two-dimension DOA estimation in massive MIMO systems. *IEEE Access*, *5*, 14020–14027.

Engineering Innovation and Design – Lam et al. (Eds)
© 2019 Taylor & Francis Group, London, ISBN 978-0-367-02959-3

Blind iterative CFO estimation for uplink interleaved OFDMA

Ann-Chen Chang, Chih-Chang Shen & Yu-Guo Zhang
Department of Information Technology, Ling Tung University, Taichung, Taiwan

ABSTRACT: This paper deals with blind iterative Carrier Frequency Offsets (CFO) estimation based on modify Propagator Method (PM) for interleaved Orthogonal Frequency Division Multiple Access (OFDMA) uplink systems, without using specific training sequences. It has been shown that the PM based on the structural algebraic properties of the received data matrix has low computational load and is only efficient in nonnoisy environments. In conjunction with the usefulness of QR/LU factorizations, the presented searching-based CFO estimators can alleviate the finite sample and noise effects. Meanwhile, an approach is suggested which finds a new CFO vector based on Taylor series expansion of one initially given. The problem of finding the new CFO vector is formulated as the closed form of a generalized eigenvalue problem, which allows one to readily solve it. Since raising the accuracy of residual CFO estimation can provide more accurate residual CFO compensation, this paper also presents an iterative estimate approach with less computational load to improve CFO estimation performance. However, the proposed estimators are not only can estimate CFO, but also have less computational load. Finally, several computer simulation results are provided for illustrating the effectiveness of the blind estimate approaches under single OFDMA data block.

1 INTRODUCTION

The increasing demand for multimedia communications with variable data rates and different quality of service requirements has recently led to a strong interest in orthogonal frequency division multiple access (OFDMA) (Miao, et al., 2009). Such systems rely crucially on carrier frequency synchronization. Since multiple user signals are mixed together in both time and frequency domains in the uplink, carrier frequency offsets (CFO) between the user transmitters and the uplink receiver may destroy the orthogonality among subcarriers, which will introduce intercarrier interference from the user itself and multiple access interference from other users. CFO estimation of such system is a challenging multi-parameter estimation problem. It is crucial to accurately estimate and compensate CFO for reliable OFDMA transmissions. So this requires an accurate CFO estimation method which becomes a more crucial task for reliable detection of transmitted data.

Some blind CFO estimators offer very high estimation performance without pilot symbols in OFD-MA systems. The general techniques for blind CFO estimation of interleaved uplink OFDMA are subspace methods such as multiple signal classification (MUSIC) (Miao, et al., 2009) and estimation of signal parameters via rotational invariance technique (ESPRIT) (Lee, et al., 2007). But these algorithms employ either eigenvalue decomposition (EVD) of data correlation matrix. Using these techniques, the computational complexity is costly and high so that these algorithms suffer from limited application in real-time signal processing environments. For minimum variance distortion less response (MVDR) (Chang & Yan, 2013), the neighboring peaks often cannot be distinguished under relatively low signal-to-noise ratio (SNR) and large active users. This implies that one user's original peak location is pulled into an adjacent user's range, severely distorting original peak location (Miao, et al., 2009). In Marcos, et al. (1995), propagator method (PM) estimates parameters without EVD of data correlation matrix. Indeed, the propagator can be estimated from data cross-correlation matrix with a least squares (LS) process. Therefore, the computational complexity of PM is lower than that of MUSIC and ESPRIT. Though computation load decreased, CFO estimation accuracy decayed because of PM is a suboptimal which does not require any modeling of background noise. For searching-based approach, complexity and estimation accuracy strictly depend on the grid size used during search. However, it is time consuming and search grid is not clear.

The goal of this paper is proceed from effectiveness and computational complexity of estimator's objective function and spectrum search. For PM, the existing structure of interleaved OFDMA system can be used to form a propagator to explore the presence of CFO. The propagator is a linear operator,

which only depends on CFO vectors and which can be easily extracted from direct data set. First, using algebraic properties of PM, cross-correlation matrix is replaced by upper triangular matrix obtained from LU factorization (Pan, 2000) or QR factorization (Chan, 1987) of cross-correlation matrix. The upper triangular matrices are partitioned to estimate propagator. Then, this paper also proposes two modify PM versions, which are referred to as U-PM and R-PM, respectively. Second, this paper also presents a blind CFO approach with Taylor series expansion of the presumed CFO vector to alleviate the above-mentioned problems in presence of CFO. Our objective is to present a first-order Taylor series expansion estimation method, which solves generalized eigenvalue decomposition (GEVD) problem to find a new CFO vector. Thus, with an iterative searching approach, the proposed iterative U-PM and R-PM (IU-PM and IR-PM) can efficiently reduce searching load and improve CFO estimate accuracy. Finally, simulation results are provided to demonstrate the effectiveness of the proposed estimators under single data block CFO estimation.

2 PRELIMINARIES

2.1 *Signal model*

Consider a M-user interleaved OFDMA uplink system with N subcarriers simultaneously communicate with the base station (BS) through an independent multipath channel. The N subcarriers are interleaved into $Q\,(Q > M)$ subchannels, each of which has $P = N/Q$ subcarriers. For convince, each subcarrier is exclusively used by only one user. Subchannel $q\,(q = 0, 1,\ldots, Q-1)$ contains subcarriers with index $\{q, Q+q, \ldots, (P-1)Q+q\}$. After cyclic prefix is removed, the received signal is $r_m(n) = \sum_{p=0}^{P-1} X_m(p) H_m(p) \exp\{\frac{j2\pi n(pQ+q_m+\varepsilon_m)}{N}\}$, where $0 \le n \le N-1$ and $\varepsilon_m \in (-0.5, 0.5)$ denotes the mth user's CFO normalized by the subcarrier spacing $2\pi/N$. $X_m(p)$ is a set of P data streams of the mth user and $H_m(p)$ represents the samples from $\bar{H}_m(k)$ at $k = pQ + q_m$. The channel frequency response $\bar{H}_m(k) = \sum_{l=0}^{L_m-1} h_m(l) \exp\{\frac{-j2\pi lk}{N}\}$ and L_m is the channel order. Let $r_m(n) = \sum_{p=0}^{P-1} X_m(p) H_m(p) \exp\{\frac{j2\pi np}{P}\} \times \exp\{\frac{j2\pi n\theta_m}{P}\}$, then the received signal $y(n) = \sum_{m=1}^{M} r_m(n) + z(n)$. $\theta_m = (q_m + \varepsilon_m)/Q$ denotes effective CFO of the mth user. $z(n)$ is the additive white Gaussian noise with zero mean and variance σ_n^2. In one OFDMA block, $\{y(n)\}_{n=0}^{N-1}$ can be arranged into a $Q \times P$ matrix (Lee, et al., 2007) and is given by

$$\mathbf{Y} = \mathbf{A}(\theta)\mathbf{S} + \mathbf{Z} \tag{1}$$

where $\mathbf{A}(\theta) = [\mathbf{a}(\theta_1)\, \mathbf{a}(\theta_2) \cdots \mathbf{a}(\theta_M)]$ with $\mathbf{a}(\theta_m) = [1\ e^{j2\pi\theta_m} \cdots e^{j2\pi(Q-1)\theta_m}]^T$ and $(\bullet)^T$ is transpose. $\mathbf{S} = \mathbf{D} \otimes (\mathbf{BW})$ with a $P \times P$ inverse discrete Fourier transform matrix $\mathbf{W}$, and $\otimes$ indicates an element-by-element product. $\mathbf{B} = [\mathbf{b}_1\ \mathbf{b}_2 \ldots \mathbf{b}_M]^T$ with $\mathbf{b}_m = [X_m(0)H_m(0)\ X_m(1)H_m(1) \ldots X_m(P-1)\ H_m(P-1)]^T$ and $\mathbf{D} = [\mathbf{d}_1\ \mathbf{d}_2 \cdots \mathbf{d}_M]^T$ with $\mathbf{d}_m = [1\ e^{j2\pi\theta_m/P} \cdots e^{j2\pi(P-1)\theta_m/P}]^T$. $\mathbf{Z}$ is a $Q \times P$ noise matrix. The ensemble correlation matrix $\mathbf{R} = E\{\mathbf{YY}^H\} = \mathbf{A}(\theta)\mathbf{R}_s\mathbf{A}^H(\theta) + \sigma_n^2\mathbf{I}_Q$, where $E\{\bullet\}$ and $(\bullet)^H$ are expectation and Hermitian transpose operation, respectively. $\mathbf{R}_s = E\{\mathbf{SS}^H\}$ and $\mathbf{I}_Q$ is a $Q \times Q$ identity matrix. Let one OFDMA block be taken, the sample correlation matrix $\hat{\mathbf{R}} = (\mathbf{YY}^H)/P$.

2.2 *Searching-based PM estimator*

The propagator of PM is a linear operator based upon a partition of CFO vectors for estimating CFO. First, the $Q \times M$ matrix $\mathbf{A}(\theta)$ with appropriate partition is defined as $\mathbf{A}(\theta) = [\mathbf{A}_1^T(\theta)\ \mathbf{A}_2^T(\theta)]^T$, where $\mathbf{A}_1(\theta)$ consists of first M rows of $\mathbf{A}(\theta)$ and $\mathbf{A}_2(\theta)$ consists of last $Q-M$ rows of $\mathbf{A}(\theta)$. The relationship between $\mathbf{A}_1(\theta)$ and $\mathbf{A}_2(\theta)$ is $\mathbf{A}_2(\theta) = \Pi^H\mathbf{A}_1(\theta)$, where transformation matrix $\Pi = [\mathbf{A}_2(\theta)\mathbf{A}_1^{-1}(\theta)]^H$ with size $M \times (Q-M)$. Define $\Phi = [\Pi^T - \mathbf{I}_{Q-M}]^T$, then $\Phi^H\mathbf{A}(\theta) = \mathbf{0}$. It shows that PM is based on noise subspace spanned by columns of matrix $\mathbf{\Phi}$. The computation of matrix $\mathbf{\Phi}$ requires a prior knowledge of active users' CFO. In practice, these CFO are unknown. However, we can use the received signal to estimate $\mathbf{\Phi}$. In the absence of noise, then $\mathbf{R}$ is partitioned as four block matrices. Let $\mathbf{R}_{11} = \mathbf{A}_1(\theta)\mathbf{R}_s\mathbf{A}_1^H(\theta)$ and $\mathbf{R}_{12} = \mathbf{A}_1(\theta)\mathbf{R}_s\mathbf{A}_2^H(\theta)$, which are $M \times M$ and $M \times (Q-M)$ matrices, respectively. There is linear relationship between $\mathbf{R}_{11}$ and $\mathbf{R}_{12}$, and $\mathbf{R}_{12} = \mathbf{R}_{11}\Pi$. The transformation matrix $\mathbf{\Pi}$ can be estimated using LS method. This is achieved by minimizing $\| \mathbf{R}_{12} - \mathbf{R}_{11}\Pi \|^2$, where denotes Euclidean norm. Then, the estimated $\hat{\Pi} = \mathbf{R}_{11}^{+}\mathbf{R}_{12}$, where $(\bullet)^+$ is pseudo inverse operation and $\mathbf{R}_{11}^{+} = (\mathbf{R}_{11}^H\mathbf{R}_{11})^{-1}\mathbf{R}_{11}^H$. Therefore, $\mathbf{\Phi} = \left[\ \hat{\Pi}^T\ -\mathbf{I}_{Q-M}\ \right]^T$ can be obtained. Due to $\mathbf{\Phi}$ is orthogonal to $\mathbf{A}(\theta)$ such that $\Phi^H\mathbf{A}(\theta) = \mathbf{0}$, exploit this to develop CFO estimator. Constructing projection matrix $\mathbf{P}_\Phi = \hat{\mathbf{\Phi}}\hat{\mathbf{\Phi}}^+$ and $\mathbf{P}_\Phi^H\mathbf{P}_\Phi = \hat{\mathbf{\Phi}}\hat{\mathbf{\Phi}}^+$. Thus, θ_m, $m = 1, 2, \cdots, M$, may be found by solving $\mathbf{P}_\Phi^H\mathbf{a}(\theta) = \mathbf{0}$ for θ. Easier to solve $\| \mathbf{P}_\Phi\mathbf{a}(\theta) \|^2 = \mathbf{a}^H(\theta)\mathbf{P}_\Phi\mathbf{a}(\theta) = 0$ for θ. Equivalently, the objective function of PM with searching grid μ_1 is given by

$$F_{\mathrm{PM}}(\theta) = \underset{\theta}{\mathrm{Max}}\ [\mathbf{a}^H(\theta)\mathbf{P}_\Phi\mathbf{a}(\theta)]^{-1} \tag{2}$$

where $\mathbf{a}(\theta)$ is the scanning vector. When $\mathbf{a}(\theta)$ is orthogonal to $\mathbf{P}_\Phi$, the denominator of $F_{\mathrm{PM}}(\theta)$

is zero, $F_{\mathrm{PM}}(\theta)$ trends to be infinite. Under the assumption $\varepsilon_m \in (-0.5, 0.5)$, we know that if a user is assigned the subchannel q, the range of its effective CFO is $((q_m - 0.5)/Q,\ (q_m + 0.5)/Q)$. Because each user lies in its own range, one-to-one mapping is possible between θ_m and ε_m. Then the normalized CFO can be calculated by $\hat{\varepsilon}_m = Q\hat{\theta}_m - q_m$ for $m = 1, 2, \cdots, M$. It is note that $\hat{\Pi} = \mathbf{R}_{11}^{+}\mathbf{R}_{12}$ is no longer valid in the presence of noise. In practice, the data are generally impaired and SNR value is not always high. Then, the performance of PM depends on signal information contained in $\mathbf{R}_{11}$ with respect to noise and its linear dependency with $\mathbf{R}_{12}$.

3 PROPOSED ESTIMATORS

3.1 *Proposed U-PM and R-PM estimators*

In this subsection, we insert a LU or QR decomposition step in PM to improve robustness to noise of PM. The properties of upper triangular matrix are used to minimize influence of model errors. Assume that $\mathbf{R}$ bears LU factorization, then it can be expressed as

$$\begin{aligned}\mathbf{R} = \mathbf{LU} &= \begin{bmatrix} \mathbf{L}_{11} & \mathbf{0} \\ \mathbf{L}_{21} & \mathbf{I}_{Q-M} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{0} & \mathbf{U}_{22} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{L}_{11}\mathbf{U}_{11} & \mathbf{L}_{11}\mathbf{U}_{12} \\ \mathbf{L}_{21}\mathbf{U}_{11} & \mathbf{L}_{21}\mathbf{U}_{12} + \mathbf{U}_{22} \end{bmatrix}\end{aligned} \tag{3}$$

Using $\mathbf{R}_{12} = \mathbf{R}_{11}\Pi$ and (3), we have $\mathbf{L}_{11}\mathbf{U}_{12} = \mathbf{L}_{11}\mathbf{U}_{11}\Pi_U$. Finally, the estimate of propagator operator using LU factorization is $\Pi_U = \mathbf{U}_{11}^{-1}\mathbf{U}_{12}$. If we calculate the following product,

$$\begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{0} & \mathbf{U}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{11}^{-1}\mathbf{U}_{12} \\ -\mathbf{I}_{Q-M} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\mathbf{U}_{22} \end{bmatrix} \tag{4}$$

It shows that columns of matrix $[(\mathbf{U}_{11}^{-1}\mathbf{U}_{12})^T\ -\mathbf{I}_{Q-M}]^T$ form a basis for eigenvectors associated with smallest eigenvalues and $\mathbf{U}_{22}$ contains smallest eigenvalues of $\mathbf{R}$. This result confirms that propagator $\left(\hat{\Pi} = \mathbf{R}_{11}^{+}\mathbf{R}_{12}\right)$ estimated from LU factorized matrix $\left(\Pi_U = \mathbf{U}_{11}^{-1}\mathbf{U}_{12}\right)$ is in accordance with propagator principle. From (4), it is showed that smallest eigenvalues are in lower-right corner of $\mathbf{U}$, that is, $\mathbf{U}_{22}$. The useful signal components are concentrated in $\mathbf{U}_{11}$ and $\mathbf{U}_{12}$. This yields a better robustness to noise compared to PM. Following similar calculations with QR factorization, $\mathbf{R}$ can be expressed as

$$\mathbf{R} = \mathbf{Q}\bar{\mathbf{R}} = \mathbf{Q}\begin{bmatrix} \bar{\mathbf{R}}_{11} & \bar{\mathbf{R}}_{12} \\ \mathbf{0} & \bar{\mathbf{R}}_{22} \end{bmatrix} \tag{5}$$

Then, we obtain $\Pi_R = \bar{\mathbf{R}}_{11}^{-1}\bar{\mathbf{R}}_{12}$. As in LU factorization, the columns of matrix $[(\bar{\mathbf{R}}_{11}^{-1}\bar{\mathbf{R}}_{12})^T\ -\mathbf{I}_{Q-M}]^T$ form a basis for eigenvectors associated with smallest eigenvalues and $\bar{\mathbf{R}}_{22}$ contains smallest eigenvalues of $\mathbf{R}$.

Let $\Phi_U = [\Pi_U^T - \mathbf{I}_{Q-M}]^T$ and $\Phi_R = [\Pi_R^T - \mathbf{I}_{Q-M}]^T$. The column vectors of Φ_U and Φ_R cannot form an orthonormal basis, we use orthogonal projection to this subspace to satisfy orthonormality and improve performance through base of orthonormality and zero-kernel space. Therefore, the projection matrices based on U-PM and R-PM are $\mathbf{P}_U = \Phi_U(\Phi_U^H\Phi_U)^{-1}\Phi_U^H$ and $\mathbf{P}_R = \Phi_R(\Phi_R^H\Phi_R)^{-1}\Phi_R^H$, respectively. It follows that CFO of users are given by positions of maxima of following functions $F_{\text{U-PM}}(\theta) = [\mathbf{a}^H(\theta)\mathbf{P}_U\mathbf{a}(\theta)]^{-1}$ and $F_{\text{R-PM}}(\theta) = [\mathbf{a}^H(\theta)\mathbf{P}_R\mathbf{a}(\theta)]^{-1}$. However, the main advantage of U-PM and R-PM is their low computational load for computing objective functions. Indeed, these methods do not require the costly EVD of correlation matrix.

3.2 *Proposed iterative estimators*

This subsection describes an iterative method to estimate CFO with less computational complexity. Assume that a CFO vector $\mathbf{a}(\theta)$ is infinitely differentiable in neighborhoods of $\hat{\theta}_m$. Let $\mathbf{a}^{(k)}(\theta_m) = \partial^k\mathbf{a}(\theta)/\partial\theta^k\,|_{\theta=\theta_m}$ for $k = 0, 1, \cdots, K$. It can be represented through Taylor series expansion at $\hat{\theta}_m$ as $\mathbf{a}(\theta) = \sum_{k=0}^{\infty} c_k(\theta)\mathbf{a}^{(k)}(\hat{\theta}_m)$, where $c_k(\theta) = (\theta - \hat{\theta}_m)^k / k!$. Using the first K terms of the series, the true CFO vector $\mathbf{a}(\theta_m)$ may be approximated as

$$\mathbf{a}(\theta_m) \cong \sum_{k=0}^{K-1} \mathbf{c}_k(\theta_m)\mathbf{a}^{(k)}(\hat{\theta}_m) = \mathbf{B}_m\mathbf{c}_m \tag{6}$$

where $\mathbf{c}_m = [c_0(\theta_m)\, c_1(\theta_m) \cdots c_{K-1}(\theta_m)]^T$ and $\mathbf{B}_m = \left[\mathbf{a}(\hat{\theta}_m)\ \mathbf{a}^{(1)}(\hat{\theta}_m) \cdots \mathbf{a}^{(K-1)}(\hat{\theta}_m)\right]$. As approaches, the approximation error $\|\mathbf{B}_m\mathbf{c}_m - \mathbf{a}(\theta_m)\|$ tends to zero. When $|\theta_m - \hat{\theta}_m|$ is small, with $K \geq 2$ would be closer to $\mathbf{a}(\theta_m)$ than, which is identical to with $K = 1$. Therefore, it is expected that better performance would be obtained if $\mathbf{B}_m\mathbf{c}_m$ could be used in place of $\mathbf{a}(\hat{\theta}_m)$. However, is unknown, though $\mathbf{B}_m$ is known. Based on the approximation given in (6), we will find a new CFO vector with the form $\mathbf{B}_m\mathbf{u}_m$, where $\mathbf{u}_m$ is a K-dimensional vector and $K < Q$. Let $\mathbf{P}$ represent $\mathbf{P}_U$ for U-PM and $\mathbf{P}_R$ for R-PM. When using $\mathbf{B}_m\mathbf{u}_m$ as the CFO vector, the output power of estimator for the mth user is $P_m = (\mathbf{u}_m^H\mathbf{B}_m^H\mathbf{P}\mathbf{B}_m\mathbf{u}_m)^{-1}$. Next, $\mathbf{u}_m$ is selected such that the estimator output power is maximized under the unit norm constraint on $\mathbf{B}_m\mathbf{u}_m$ (Choi, 2011):

$$\underset{\mathbf{u}}{\text{Min}}\ \mathbf{u}_m^H \mathbf{B}_m^H \mathbf{P} \mathbf{B}_m \mathbf{u}_m \text{ subject to } \| \mathbf{B}_m \mathbf{u}_m \|^2 = 1 \tag{7}$$

Introducing Lagrange multiplier method, (7) is transformed into unconstrained minimization as $\underset{\mathbf{u}}{\text{Min}}\ f(\gamma_m, \mathbf{u}_m)$, where $f(\gamma_m, \mathbf{u}_m) = \mathbf{u}_m^H \mathbf{B}_m^H \mathbf{P}_\Phi \mathbf{B}_m \mathbf{u}_m + \gamma_m (1 - \mathbf{u}_m^H \mathbf{B}_m^H \mathbf{B}_m \mathbf{u}_m)$ and γ_m is Lagrange multiplier. Taking derivative of $f(\gamma_m, \mathbf{u}_m)$ with respect to $\mathbf{u}_m$ and setting it to zero, we have

$$\mathbf{B}_m^H \mathbf{P} \mathbf{B}_m \mathbf{u}_m = \gamma_m \mathbf{B}_m^H \mathbf{B}_m \mathbf{u}_m \tag{8}$$

Recalling that $\| \mathbf{B}_m \mathbf{u}_m \|^2 = 1$ and substituting (8) into $P_m = (\mathbf{u}_m^H \mathbf{B}_m^H \mathbf{P} \mathbf{B}_m \mathbf{u}_m)^{-1}$, it follows that $P_m = \gamma_m^{-1}$. Hence, solution $\overline{\mathbf{v}}_m$ to (7) corresponds to eigenvector associated with smallest eigenvalue of GEVD problem of (8). The new CFO scanning vector $\overline{\mathbf{v}}_m$ is then given by $\overline{\mathbf{v}}_m = \mathbf{B}_m \mathbf{u}_m$. However, owing to the finite samples effect, $\overline{\mathbf{v}}_m$ is not always proportional to $\mathbf{a}(\theta_m)$. To make the first entry equal to one for the actual CFO vector, we define a vector shown as $\mathbf{v}_m = (\overline{v}_{m,1})^{-1} \overline{\mathbf{v}}_m$, where $\overline{v}_{m,1}$ denotes first entry of $\overline{\mathbf{v}}_m$. Since the amplitude of each entry of $\mathbf{a}(\theta_m)$ is equal to one, an appropriate estimate $\mathbf{a}(\hat{\theta}_m)$ can then be obtained from $\hat{a}_{m,p} = v_{m,p} / |v_{m,p}|$, $p = 1, 2, \ldots, Q$, and $v_{m,p}$ is the pth entry of $\mathbf{v}_m$. For mth user, the CFO vector is $\mathbf{a}(\hat{\theta}_m) = \left[1\ e^{j2\pi\hat{\theta}_m} \cdots e^{j2\pi(Q-1)\hat{\theta}_m}\right]^T$. $\hat{\mathbf{g}}_m = angle\{\mathbf{a}(\hat{\theta}_m)\}$ can be obtained, where $angle\{\bullet\}$ is to get phase angle for each element of $\mathbf{a}(\hat{\theta}_m)$. Note that, $\hat{\mathbf{g}}_m = \left[0\ 2\pi\hat{\theta}_m \ldots 2\pi\ (Q-1)\ \hat{\theta}_m\right]^T$, and then LS principle is adopted to estimate $\hat{\theta}_m$. LS fitting is given by

$$\underset{\mathbf{k}_m}{\text{Min}} \| \mathbf{T}\mathbf{k}_m - \hat{\mathbf{g}}_m \|^2$$
$$\text{and } \mathbf{T} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 2\pi & \cdots & 2(Q-1)\pi \end{bmatrix}^T \tag{9}$$

where $\mathbf{k}_m = [k_{m0}\ \ k_{m1}]^T \in R^{2\times 1}$ is an unknown parameter vector, k_{m1} is the estimated value of θ_m, and k_{m0} is the other estimation parameter. The LS solution for $\mathbf{k}_m$ is given by

$$\hat{\mathbf{k}}_m = (\mathbf{T}^T \mathbf{T})^{-1} \mathbf{T}^T \hat{\mathbf{g}}_m \tag{10}$$

Then, θ_m is estimated via $\hat{\theta}_m = \hat{k}_{m1}$. Since each user lies in its own range, one-to-one mapping is possible between θ_m and ε_m. Hence, $\hat{\varepsilon}_m = Q\hat{\theta}_m - q_m$. However, if estimating deviation $|\theta_m - \hat{\theta}_m|$ is quite large, then it requires an iterative approach that updates estimating deviation toward to actual value. Finally, based on the above stated estimation scheme, the major steps for implementing IU-PM and IR-PM is described as follows:

Step 1. Set the initial estimates for each user.
Step 2. Update counter i and calculate $\mathbf{B}_m^i$.
Step 3. Perform GEVD on $(\mathbf{B}_m^i)^H \mathbf{P} \mathbf{B}_m^i \mathbf{u}_m^i = \gamma_m (\mathbf{B}_m^i)^H \mathbf{B}_m^i \mathbf{u}_m^i$ to obtain eigenvector $\mathbf{u}_m^i$ which corresponds to smallest eigenvalue. Then, $\overline{\mathbf{v}}_m^i = \mathbf{B}_m^i \mathbf{u}_m^i$.
Step 4. Calculate $\mathbf{v}_m^i = (\overline{v}_{m,1}^i)^{-1} \overline{\mathbf{v}}_m^i$ and update $\hat{a}_{m,p} = v_{m,p} / |v_{m,p}|$, $p = 1, 2, \ldots, Q$, to obtain $\mathbf{a}(\hat{\theta}_m^i)$.
Step 5. Calculate (10) to obtain $\hat{k}_{m1}^i$ and update $\hat{\theta}_m^i = k_{m1}^i$.
Step 6. Check convergence of $\hat{\theta}_m^i$. If $|\theta_m^i - \theta_m^{i-1}| < \mu_2$, output $\hat{\theta}_m = \hat{\theta}_m^i$. Otherwise, let $i = i + 1$ and go back to Step 2.
Step 7. The mth user's CFO estimation is $\hat{\varepsilon}_m = \hat{\theta}_m^i \times Q - q_m$.

4 SIMULATION RESULTS

The parameters in OFDMA uplink system are $M = 8$, $N = 1024$, $Q = 32$ and $P = 32$. The symbols used for CFO estimation are modulated with binary phase-shift-keying modulation and the average received signal power from all users is the same. The subcarrier allocation rule for each user is perfectly known at BS. Each user transmits signal to BS through independent multipath channel. The channel taps $h_m(l)$ are modeled as statistically independent Gaussian random variables with zero mean and an exponentially decaying power profile, $E[h_m(l)^2] = \alpha_l e^{(-l/5)}$, $0 \le l \le L_m - 1$, where α_1 is a normalized factor used to set channel power to unity (J.H. Lee, et al. 2007). $L_m = 10$ for every user in each Monte Carlo test. In simulation, we also assume that channel state will not vary in the duration of one OFDMA block. As an index of evaluation, input SNR and mean square error (MSE), were defined as $\text{SNR} = 10 \log E\{r_m(n)^2\} / \sigma_n^2$ and $\text{MSE} = (1/M\Pi) \Sigma_{\rho=1}^{\Pi} \Sigma_{m=1}^{M} (\hat{\varepsilon}_m^\rho - \varepsilon_m^\rho)^2$ with $\Pi = 10^3$ is number of runs. For all simulations, the CFO of the active users are generated as random variables uniformly distributed in (–0.5, 0.5) and mutually independent among users. Each of simulation results presented is after one OFDMA block processed with independent noise samples for each run.

First, in the scenario range [0 dB, 30 dB], the proper choice for PM, U-PM, and R-PM is $\mu_1 = 10^{-4}$ and $\mu_2 = 10^{-2}$ under $K = 2$ for IU-PM and IR-PM. Fig. 1 shows MSE versus selected terms K of Taylor series expansion for IU-PM and IR-PM with $\mu_2 = 10^{-2}$. Although $\mathbf{a}(\theta_m)$ is assumed to be infinitely differentiable in the neighborhoods of $\mathbf{a}(\hat{\theta}_m)$, which can be represented through as using the first K terms of Taylor series expansion at $\hat{\theta}_m$. In theory, increasing the terms of Taylor series expansion, the approximated CFO vector will be converged to the true CFO vector.

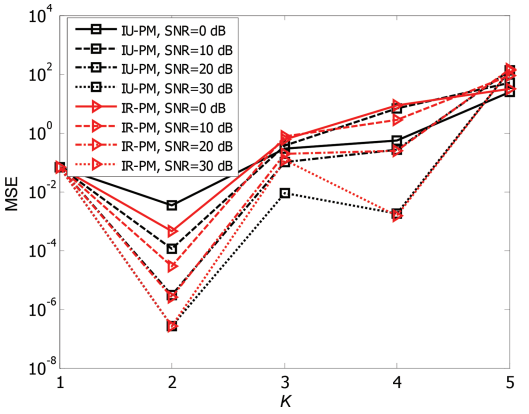

Figure 1. MSE performance of IU-PM and IR-PM versus K under $\mu_2 = 10^{-2}$.

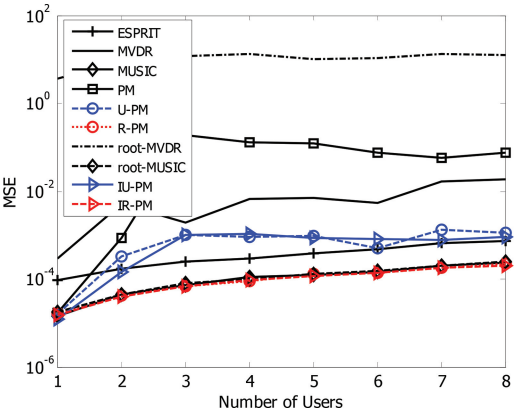

Figure 2. MSE versus number of users.

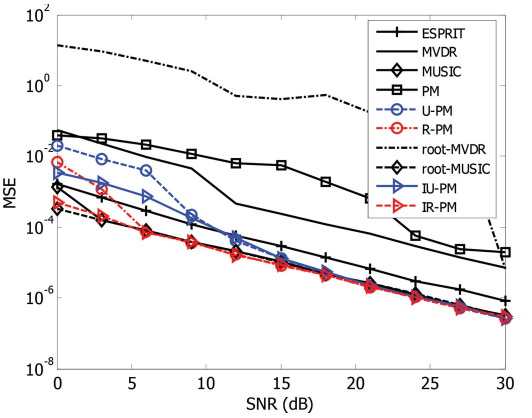

Figure 3. MSE versus SNR.

But, the effective degree of freedom of estimator will be reduced and then the resolution will be reduced. From this figure, we see that the IU-PM and IR-PM with $K = 2$ substantially outperforms one with $K = 3$, $K = 4$, or $K = 5$. Therefore, a proper choice is $K = 2$.

Second, the performance of the proposed estimators are evaluated and compared with MUSIC, root-MUSIC, MVDR, root-MVDR, and ESPRIT. Fig. 2 shows MSE versus the number of users. Assume that all active users have equal power with SNR = 3 dB and CFO are uniformly random variables distributed in the range of $(-0.5, 0.5)$. It is shown that the estimation performance of the proposed IU-PM and IR-PM are more accurate than that of MVDR and PM as active user' number increases. Fig. 3 presents MSE of CFO estimation versus SNR. It is noted that the statistical behavior of MVDR for one data block appears difficult to establish. Therefore, the performance of root-MVDR will become worse than MVDR at low SNR. Observe that the proposed IU-PM and IR-PM have very good performance to U-PM and R-PM, especially under low SNR. Again, the proposed estimators have better estimate performance than direct using PM estimator.

Next, let us compare searching efficiency of U-PM, R-PM, IU-PM, and IR-PM in practical. The averaged searching numbers were defined as $F_i = (1/M\Pi)\Sigma_{\rho=1}^{\Pi}\Sigma_{m=1}^{M}F_{i,m}^{\rho}$ for $i = 1, 2$. Assume that searching grid of U-PM and R-PM is $\mu_1 = 10^{-4}$, then number of searching $F_1 = 10001$. The terminating error value of IU-PM and IR-PM is $\mu_2 = 10^{-2}$, respectively. Especially, we find that estimation accuracy of IU-PM and IR-PM keeps still after 2 iterations, that is to say these estimators could achieve convergence with 2 iterations in scenario range [0 dB, 30 dB]. It is noted that the proposed approach is fast adaptively corrected from the previous iteration stage in spite of how well the initial guess is performed, and then it makes the convergence rate enhanced.

5 CONCLUSIONS

This paper has presented the iterative CFO estimators to increase estimate accuracy under single data block. The existing PM is a LS solution and very sensitive to noise. First, it demonstrates the usefulness of QR and LU factorizations of data correlation matrix are efficient in noisy situations. In order to reduce computational load of search and exhibit good performance, the CFO estimators based on the iterative search approach are also developed. In particular, the iterative search problem of finding the new CFO vector is formulated as closed form of a GEVD problem, which allows one to readily solve it and does not require a strict initial requirement setting. It not only can estimate CFO, but also has less computational load.

ACKNOWLEDGMENT

This research was supported by the Ministry of Science and Technology under grant number MOST 107-2221-E-275-005.

REFERENCES

Chan, T.F., 1987. *Linear Algebra and Its Applications,* 88–89 67–82.

Chang, A.C., S.H. Yang, 2013. *Computing, Communications and IT Applications Conference,* 153–158.

Choi, Y.H., 2011. *Electronics Letters,* 47(15), 840–841.

Lee, J.H., S. Lee, K.J. Bang, 2007. *IEEE Trans. Vehicular Technology,* 56(5), 3227–3231.

Marcos, S., A. Marsal, M. Benidir, 1995. *Signal Processing,* 42(2), 121–138.

Miao, R., J. Xiong, L. Gui, et al., 2009. *IEEE Trans. Consumer Electronics,* 55(3), 1039–1044.

Pan, C.T., 2000. *Linear Algebra and Its Applications,* 316(1–3), 199–222.

Engineering Innovation and Design – Lam et al. (Eds)
© 2019 Taylor & Francis Group, London, ISBN 978-0-367-02959-3

Ultra-broadband dispersion-flattened/polarization-maintaining photonic-crystal fiber

Yi-Jun Cai, Shu-Han Chu & Jui-Ming Hsu
Department of Electro-Optical Engineering, National United University, Miaoli, Taiwan

ABSTRACT: We designed a dual-function nearly-zero Dispersion-Flattened/Polarization-Maintaining Photonic-Crystal Fiber (DF/PM PCF) for use with ultra broadband. The numeric results show that the DF/PM PCF delivers nearly-zero dispersion values between –1.00 and –0.12 ps/nm/km over a bandwidth range of 482 nm (1218–1700 nm). The operational band ranges from the O band (1260–1360 nm) to the U band (1625–1675 nm). In addition, the birefringence values of the proposed DF/PM PCF were estimated and shown to be between 0.93×10^{-3} and 1.25×10^{-3} in the operational band. Confinement losses for some wavelengths in the operational band were calculated and confirmed to be reasonable. The numeric results indicate that the DF/PM PCF has confinement losses of 2.75×10^{-3} to 3.39×10^{-1} dB/km in the operational band. Furthermore, the proposed DF/PM PCF was simulated and confirmed to work in single mode in the operational band. The device can be well used in a Dense Wavelength Division Multiplexing (DWDM) communication system.

1 INTRODUCTION

In traditional optical fiber communication systems, Dispersion-Flattened Fiber (DFF) may effectively reduce the chromatic dispersion and increase transmission rate over the broadband wavelength range. For a traditional step-index DFF, it is possible to use multiple cladding layers and tailor the index distribution such that the chromatic dispersion is relatively small over a wide wavelength range (Cohen et al., 1982; Jang et al., 1982; Bhagavatula et al., 1983).

On the other hand, Polarization-Maintaining Fiber (PMF) can reduce the governing factor of small random birefringence fluctuations in the fiber significantly, and restrain the Polarization Modal Dispersion (PMD). Conventional Modified Chemical Vapor Deposition (MCVD) PMFs (such as elliptical-core fibers, bow-tie-structured fibers, and PANDA fibers) typically have a modal birefringence with a value of about 10^{-4} order.

Photonic-Crystal Fibers (PCFs) have attracted substantial attention recently because of their outstanding properties and great design flexibility. Several designs for DFFs and PMFs based on PCF have been proposed, such as Dispersion-Flattened Photonic-Crystal Fibers (DFPCFs) (Wu & Chao, 2005; Hou et al., 2014; Gundu et al., 2006; Hsu, 2016) and Polarization-Maintaining Photonic-Crystal Fibers (PMPCFs) (Hansen et al., 2001; Suzuki et al., 2001; Hsu et al., 2012, 2013; Nasilowski et al., 2005; Chen et al., 2008; Ortigosa-Blanch et al., 2000; Hsu, 2017).

In order to reduce excessive loss and poor matching between DFPCF and PMPCF cascaded in a system, this study merged the functions of DFF and PMF in one PCF, and designed a dual function of Dispersion-Flattened and Polarization-Maintaining Photonic-Crystal Fiber (DF/PM PCF). This study used the dispersion slope matching method to design a PCF with nearly-zero flattened dispersion and polarization-maintaining characteristics in the broadband range.

2 PRINCIPLES

2.1 *Birefringence*

Birefringence is defined as the difference between effective refractive indices of two split degeneration-pair fundamental modes, and can be written as:

$$B = \left| n_x - n_y \right| \tag{1}$$

where n_x and n_y are the effective indices of the x- and y-polarization modes, respectively.

2.2 *Chromatic dispersion*

The chromatic dispersion coefficient D, material dispersion coefficient D_M, and waveguide dispersion coefficient D_G, are defined as:

$$D(\lambda) = D_M(\lambda) + D_G(\lambda) = \left(\frac{-\lambda}{c} \right) \frac{d^2 n_{eff}}{d\lambda^2} \tag{2}$$

$$D_M(\lambda) = \left(\frac{-\lambda}{c}\right)\frac{d^2 n_m}{d\lambda^2} \quad (3)$$

$$D_G(\lambda) = \left(\frac{-\lambda}{c}\right)\frac{d^2 n_g}{d\lambda^2} \quad (4)$$

where c is the speed of light in a vacuum, n_{eff} is the effective index of the fundamental guided mode in a fiber, n_m is the refractive index of the material (silica in this case), and the effective index of the waveguide n_g is the refractive index evaluated by considering the condition of the material refractive index to be wavelength-independent.

The chromatic dispersion D is a sum of the material dispersion D_M and the waveguide dispersion D_G, as shown in Equation 2. For ideal conditions at wavelength λ, zero chromatic dispersion must be achieved, that is, $D(\lambda) = D_M(\lambda) + D_G(\lambda) = 0$. Therefore, to achieve a nearly-zero ultra-flattened dispersion PCF, one can tailor the waveguide dispersion $D_G(\lambda)$ to fit in with:

$$D_G(\lambda) \approx -D_M(\lambda) \quad (5)$$

over a bandwidth range as wide as possible.

2.3 *Confinement Losses (L_C)*

The confinement loss (L_C) for a mode propagated in a PCF can be estimated by the imaginary part of the effective index (n_{eff}):

$$L_C = 8.686 \times k_o \times Im[n_{eff}] \quad (6)$$

where k_o is the wave number in free space, and $Im[n_{eff}]$ represents the imaginary part of the n_{eff}.

3 SIMULATION

The cross-sectional structure of the proposed DF/PM PCF is shown in Figure 1. Considering ease of fabrication, we use an index-guiding PCF with a triangular lattice of air holes with a pitch of Λ to design the DF/PM PCF. The background material is undoped silica; the refractive index of silica can be calculated using the Sellmeier equation (Malitson, 1965):

$$n_m(\lambda) = \left[1 + \frac{A_1\lambda^2}{\lambda^2 - (\lambda_1)^2} + \frac{A_2\lambda^2}{\lambda^2 - (\lambda_2)^2} + \frac{A_3\lambda^2}{\lambda^2 - (\lambda_3)^2}\right]^{1/2} \quad (7)$$

where λ is the operating wavelength. The Sellmeier constants A_1, A_2, A_3, λ_1, λ_2 and λ_3 are 0.6961663, 0.4079426, 0.8974794, 0.0684043, 0.1162414 and 9.896161, respectively.

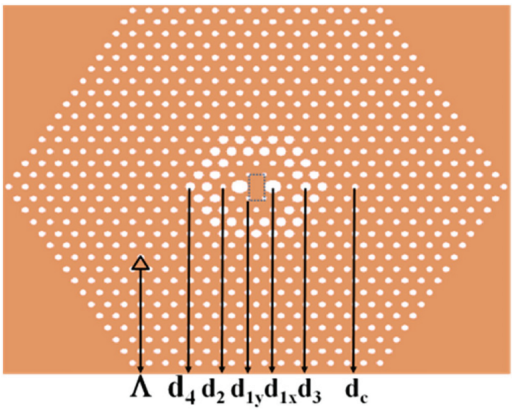

Figure 1. Cross-sectional structure of the proposed DF/PM PCF.

In Figure 1, the diameters of the air holes of the first to fourth layers are denoted as d_{1x}, d_{1y}, d_2, d_3 and d_4, respectively. To enlarge B value for polarization maintenance, one must destroy the six-fold symmetry of the fiber. Therefore, holes with a larger d_{1x} diameter and a smaller d_{1y} diameter are introduced into the first layer. This results in a split degeneration doublet pair and enlarges the modal birefringence in the PCF. However, the values of d_{1x} and d_{1y} may influence the chromatic dispersion; therefore they must be designed in combination with the parameter design of the other layers.

To effectively reduce the confinement loss of the PCF, the inner four layers are surrounded by eleven layers of air holes, which is the cladding of the fiber, with a diameter of d_c.

On the basis of Equation 5, we trimmed the parameters Λ, d_{1x}, d_{1y}, d_2, d_3 and d_4 to tailor the $D_G(\lambda)$ curve to overlap with the $-D_M(\lambda)$ curve over a bandwidth range as wide as possible.

After the D_G-tailoring procedure, appropriate geometric parameters are selected and a DF/PM PCF with a high birefringence and near-zero dispersion over a wide band is achieved. Afterwards, the real and imaginary parts of the n_{eff} of the DF/PM PCF structure can be calculated using the Beam Propagation Method (BPM). Finally, the imaginary part of n_{eff} is substituted into Equation 6 to deduce the confinement loss.

4 RESULTS AND DISCUSSION

After the D_G-tailoring procedure, the optimum parameters are identified as follows: Λ = 1.815 μm, d_{1x} = 1.800 μm, d_{1y} = 0.495 μm, d_2 = 0.660 μm, d_3 = 1.120 μm, d_4 = 1.100 μm, and d_c = 0.750 μm. The overlap between $D_G(\lambda)$ and $-D_M(\lambda)$ curves

is shown in Figure 2. As per the figure, $D_G(\lambda) \approx -D_M(\lambda)$ between the wavelengths of 1218 and 1923 nm. This means that the waveguide dispersion nullifies the material dispersion in this wavelength range.

Figure 3 demonstrates the dependence of the chromatic dispersion ($D(\lambda) = D_M(\lambda) + D_G(\lambda)$) and birefringence on wavelength for the proposed DF/PM PCF. As shown in the figure, the DF/PM PCF delivers nearly-zero dispersion values between –1.00 and –0.12 ps/nm/km over a bandwidth range of 482 nm (1218–1700 nm). In the meantime, the birefringence values are in the range of 0.93×10^{-3} to 1.25×10^{-3} in the operational band. The birefringence values are more than ten times those of conventional PMFs.

By using Equation 6, the confinement losses of the proposed DF/PM PCF for some wavelengths in the operational band are calculated, and are shown in Figure 4. As per the figure, the confinement losses are in the range of 2.75×10^{-3} to 3.39×10^{-1} dB/km, which is reasonable for a fiber optic communication link in the operational band.

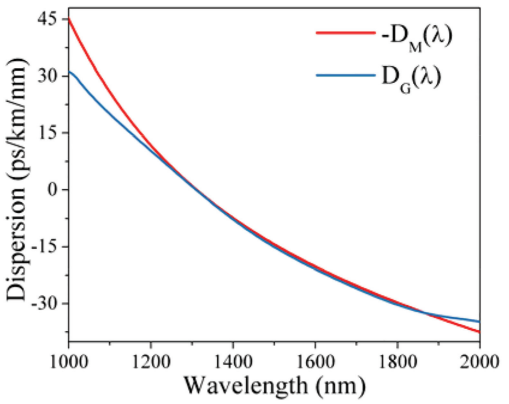

Figure 2. Tailoring the $D_G(\lambda)$ curve to overlap the $-D_M(\lambda)$ curve over a bandwidth range as wide as possible.

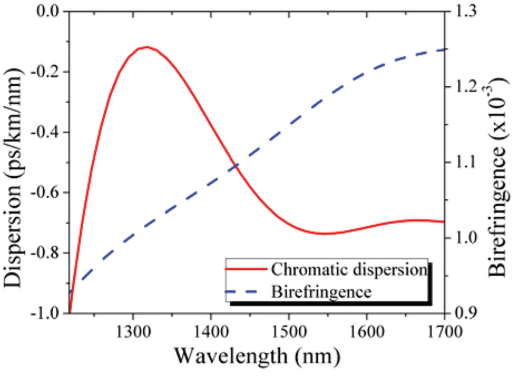

Figure 3. Dependence of the chromatic dispersion and birefringence on wavelength for the proposed DF/PM PCF.

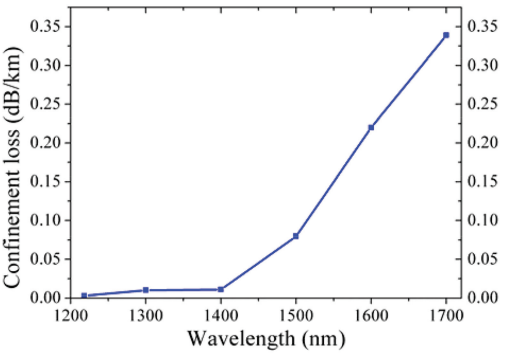

Figure 4. Relationship between the confinement losses of the proposed DF/PM PCF and wavelengths in the operational band.

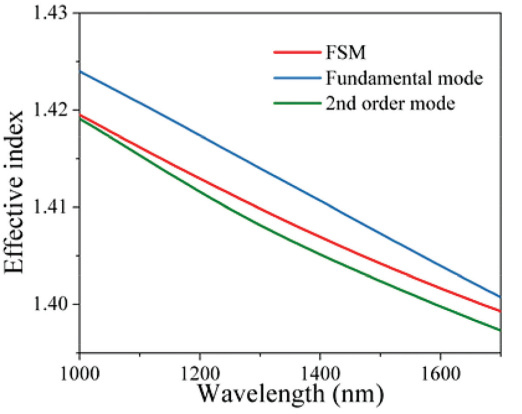

Figure 5. Relationship of effective indices between the fundamental mode, second-order mode, and FSM of the DF/PM PCF in the operational band.

The fiber must work in single mode; otherwise, a more serious modal dispersion will be excited. The relationship of effective indices between the fundamental mode, second-order mode, and Fundamental Space-filling Mode (FSM) of the proposed DF/PM PCF are displayed in Figure 5. The effective index of the fundamental mode, second-order mode, and FSM are denoted as n_1, n_2 and n_{FSM}, respectively. The n_{FSM} represents the effective index of the cladding (Birks et al., 1997). Therefore, the condition of single-mode operation of the fiber is (λ)

$$n_1(\lambda) > n_{FSM}(\lambda) > n_2(\lambda) \tag{8}$$

As per Figure 5, the effective indices of the fundamental mode, second-order mode, and FSM of the proposed DF/PM PCF well-fulfill the single-mode condition (Equation 8) for the whole operational band.

5 CONCLUSION

In this paper, a dual-function nearly-zero dispersion-flattened and polarization-maintaining photonic-crystal fiber is presented. Over an operational wavelength range of 1218–1700 nm, the chromatic dispersion fluctuates marginally from –1.00 to –0.12 ps/nm/km. The birefringence of the device varies from 0.93×10^{-3} to 1.25×10^{-3}, which is more than ten times that of conventional PMFs. The confinement losses are in the range of 2.75×10^{-3} to 3.39×10^{-1} dB/km, which is reasonable for fiber optic communication links. Furthermore, the proposed DF/PM PCF is confirmed to work in single mode over the entire operational band. The bandwidth of the operational band is 482 nm, providing for a wide range from O band to U band. The device is especially suitable for a DWDM communication system.

REFERENCES

Bhagavatula, V.A., Spotz, M.S., Spotz, W.F., Love, D.B. Keck, (1983). "Segmented-core single-mode fibers with low loss and low dispersion," *Electron. Lett.*, *19*, 317–318.

Birks, T.A., J.C. Knight, & P. St. J. Russell, (1997). "Endlessly single-mode photonic crystal fiber," *Opt. Lett.*, *22*, 961–963.

Chen, M., S.G. Yang, F.F. Yin, H.W. Chen, S.Z. XIE, (2008). "Design of a new type high birefringence photonic crystal fiber," *Optoelectron. Lett.*, *4*, 0019–0022.

Cohen, L.G., W.L. Mammel, S. Lumish, (1982). "Tailoring the shapes of dispersion spectra to control bandwidths in single-mode fibers," *Opt. Lett.*, *7*, 183–185.

Gundu, K.M., M. Kolesik, J.V. Moloney, (2006). "Ultra-flattened-dispersion selectively liquid-filled photonic crystal fibers," *Opt. Express*, *14*, 6870–6878.

Hansen, T.P., J. Broeng, S.E.B. Libori, E. Knudsen, A. Bjarklev, J.R. Jensen, H. Simonsenet, (2001). "Highly birefringent index-guiding photonic crystal fibers," *IEEE Photon. Technol. Lett.*, *13*, 588–590.

Hou, J., J. Zhao, C. Yang, Z. Zhong, Y. Gao, S. Chen, (2014). "Engineering ultra-flattened-dispersion photonic crystal fibers with uniform holes by rotations of inner rings," *Photonics Res.*, *2*, 59–63.

Hsu, J.M. (2016). "Tailoring of nearly zero flattened dispersion photonic crystal fibers," *Opt. Commun.*, *361*, 104–109.

Hsu, J.M. (2017). "Systematic design of highly birefringent photonic crystal fibers," Appl. Phys. B, 123, 73.1–73.9.

Hsu, J.M., C.L. Lee, J.S. Horng, J.J.H. Kung, (2013). "Polarization-maintaining photonic crystal fiber with rim-touched defect-holes," *Opt. Commun.*, *298–299*, 125–128.

Hsu, J.M., G.S. Ye, D.L. Ye, (2012). "Birefringence and loss consideration for a highly birefringent photonic crystal fiber," *Fiber Integrated Opt.*, *31*, 11–22.

Jang, S.J., L.G. Cohen, W.L. Mammel, M.A. Shaifi, (1982). "Experimental verification of ultra-wide bandwidth spectra in double-clad single-mode fiber," *Bell Syst. Tech. J.*, *61*, 385–390.

Malitson, I.H. (1965). "Interspecimen comparison of the refractive index of fused silica," *J. Opt. Soc. Am*, *55*, 1205–1209.

Nasilowski, T., T. Martynkien, G. Statkiewicz, M. Szpulak, J. Olszewski, G. Golojuch, W. Urbanczyk, J. Wojcik, P. Mergo, M. Makara, F. Berghmans, H. Thienpont, (2005). "Temperature and pressure sensitivities of the highly birefringent photonic crystal fiber with core asymmetry," *Appl. Phys. B*, *81*, 325–331.

Ortigosa-Blanch, A., J.C. Knight, W.J. Wadsworth, J. Arriaga, B.J. Mangan, T.A. Birks, P. St. J. Russell, (2000). "Highly birefringent photonic crystal fibers," *Opt. Lett.*, *25*, 1325–1327.

Suzuki, K., H. Kubota, S. Kawanishi, M. Tanaka, M. Fujita, (2001). "Optical properties of a low-loss polarization-maintaining photonic crystal fiber," *Opt. Express*, *9*, 676–680.

Wu, T.L., C.H. Chao, (2005). "A novel ultraflattened dispersion photonic crystal fiber," *IEEE Photonics Technol. Lett.*, *17*, 67–69.

Computer science and information technology

Engineering Innovation and Design – Lam et al. (Eds)
© 2019 Taylor & Francis Group, London, ISBN 978-0-367-02959-3

A novel 3D-sudoku model for image steganography

Ji-Hwei Horng, Chih-Hung Lee, Hao-Ying Lu & Jhe-Huei Liao
Department of Electronic Engineering, National Quemoy University, Kinmen, Taiwan

ABSTRACT: In this paper, we propose a novel 3D-sudoku model for image steganography. Our 3D-sudoku table is constructed using a typical 9×9 Sudoku table. We propose a procedure to permute the elements in the 2D table. Then, different permutation versions are piled up to form a $9 \times 9 \times 9$ 3D-Sudoku table. Based on our permutation scheme, the resulting 3D-sudoku table can hide secret data with a very low PSNR value.

1 INTRODUCTION

The communication bandwidth and storage media have greatly progressed in recent years. Social media such as Facebook, Instagram, and Twitter provide a lot of images or videos in our daily life. Frequent uploading and downloading of images or videos make an excellent environment for applying the image steganography technique to transmit secret information.

The most common steganography schemes for image data-hiding are the modification direction method (Niu, et al., 2015) and the Sudoku-based method (Chang, et al., 2008; Sarada, et al., 2012; Usha, et al., 2015; Nguyen, et al., 2015; Xia, et al., 2016). The modification direction method has good embedding capacity. However, its encryption scheme is rather regular and its security should be enhanced by combining with some other permutation method. The Sudoku-based image steganography method utilizes the diversity of the Sudoku table to improve the security level with a trade-off of reducing the embedding capacity.

In this paper, we propose a novel 3D-Sudoku model for image steganography. We will review the related works, propose our method, then, analyze the theoretical PSNR value of our method by applying a simulation program.

2 RELATED WORKS

The Chang's method (Chang, et al., 2008) is inspired from Zhang and Wang's method and Sudoku solutions. A selected Sudoku solution is used to guide cover pixels' modification in order to imply secret data. Sarada and BalaSwamy (Sarada, et al., 2012) proposed a method that utilize a 16×16 Sudoku table to hide 4 digits secret data in the red and blue components for each image pixel. Nguyen and Chang (Nguyen, et al., 2015) devised a reversible scheme to hide data based on the Sudoku technique. Their scheme is ingenious. However, the data-hiding capacity is much lower than the original irreversible version.

Xia, et al. (2016) proposed a 3D-Sudoku model. Before introducing their model, let's first examine a typical 9×9 2D-Sudoku matrix shown in Fig. 1. It is composed of nine sub-matrices of size 3×3. Each sub-matrix contains nine distinct digits from 1 to 9. This property also holds for each row or each column. This provides three different ways for the data to hide in the Sudoku technique.

The 3D-Sudoku model proposed by Xia et al. (2016) is an $8 \times 8 \times 8$ 3D matrix, which is composed of eight 3D sub-matrices of size $4 \times 4 \times 4$. Each sub-matrix contains distinct numbers from 0 to 63. It is the same for each 8×8 plane in all directions of the X, Y, and Z axes. This provides four ways for data to hide in this model. This scheme can hide 6 bits of secret data into 3 pixels. The hiding capac-

1	5	3	9	8	4	7	6	2
8	4	2	7	3	6	1	5	9
6	9	7	5	1	2	8	3	4
2	3	8	6	7	1	4	9	5
9	7	4	3	2	5	6	1	8
5	1	6	8	4	9	3	2	7
7	6	5	4	9	3	2	8	1
3	8	1	2	5	7	9	4	6
4	2	9	1	6	8	5	7	3

Figure 1. A typical 9×9 Sudoku table from the website. http://www.nikoli.co.jp/puzzles/1/index_text-e.htm.