Prioritized Multiuser Transmission Using Virtual Diagonalization and Weighted Sum-MSE

Khaled Hassan and Werner Henkel
School of Engineering and Science, Jacobs University, Transmission Systems Group (TrSyS)
28759 Bremen, Germany, {k.hassan & w.henkel}@jacobs-university.de

Abstract— **In this paper, a double-prioritized transmission is realized in a non-orthogonal multiuser MIMO-OFDM transmission. First, different qualities-of-service (QoSs) are realized among different users via optimizing their transceiver linear filters by minimizing the weighted sum mean-squared error (MSE), i.e., using different weighting factors. These factors are selected according to the required QoSs. Second, after realizing multiuser transmission with a minimum MSE (MMSE), we opt for implementing a prioritized bit-loading to realize three different classes. The bit-loading is performed according to the achievable minimum sum MSE exploiting its virtual diagonalization among users. Hence, bits and power values are allocated according to the MMSE values instead of the SNR. Our double-prioritized adaptive modulation outperforms the non-adaptive multiuser MIMO-OFDM.**

Index Terms— **UEP, MIMO-OFDM, multiuser MIMO system, adaptive modulation, duality, and MMSE**

I. INTRODUCTION

UNLIKE the scalar AWGN broadcast channel (BC), the MIMO BC is in general non-degraded [3]. This means that the users in MIMO cases cannot be sorted according to their channel gains. A straight-forward solution for this problem was proposed in [1], known as block-diagonalization, which allows each user to transmit multiple data streams on the other users' null-spaces. This means to change the multiuser problem into an equivalent parallel (non-interfering) and also a degraded one. Hence, we exploited the degraded behavior of this algorithm in [2] to realize different QoSs by devoting arbitrary data-rates and arbitrary SER to each user using adaptive hierarchical modulation. The main drawback of this scheme is the strict constraint on the number of antennas and the requirement for an accurate channel state information (CSI).

To relief the strict constraint on the number of antennas and to design a more robust multiuser MIMO, a non-diagonal beamforming that exploits the duality between the uplink and the downlink is required. Thus, the complexity needed to find the optimum BC transmit filters have been reduced by utilizing the duality to its equivalent multiple-access channel (MAC). An early study by Boche et al. was considering an SINR constraint problem in order to realize the BC beamformers [3]. Another important type of this duality, which is the MSE-duality, was introduced in [5]. Currently, the more practical ideas are considering the traditional Lagrangian iterative dual method [6] which reduces the overall complexity. However, they can only guarantee achieving a local minimum instead of the global one [6].

In order to realize different QoSs amongst users, Schubert et al. discussed the feasibility region assuming a constrained target SINR value ($\mathrm{SINR_T}$) [3]. Thus, the system is said to be feasible if the achievable $\min \mathrm{SINR}_i$ is greater than or equal to a certain $\mathrm{SINR_T}$, i.e., $\min \mathrm{SINR}_i/\mathrm{SINR_T} \geq 1$. Using the analogy between the SINR and the MSE, the authors in [8] considered another QoS constraint, which is the minimum MSE (MMSE) of the u^{th} user. The MMSE_u in this case is given as a function of the user SINR and the transmit filter $\mathbf{F}_u$ such that

$$E_u = f\left(1/\mathrm{SINR}_u(\mathbf{F_u})\right) \ . \tag{1}$$

In other words, maximizing the SINR leads to minimizing E_u. This, in general, aims at optimizing the overall performance, e.g., maximizing the throughput. Moreover, optimizing the MSE is very suitable for switching off arbitrary users, or even data-streams, if they are facing poor channel qualities or, equivalently, very low SINR [6]. Generally speaking, the optimum resource allocation based on minimum MSE (MMSE) optimization can be handled knowing that the relation between MMSE and SINR.

Accordingly, we implement a UEP bit and power allocation (similar to [12]) to realize data with different priorities among each user's subcarriers. However, this time according to the unique MMSE of each user, i.e., using their MMSE_u, as if they are allocating non-interfering parallel channels. Furthermore, arbitrary margin separations are preserved among users by devoting different MSE weighting factors w_u.

The rest of this paper is organized as follows. Section 2 discusses the multiuser MIMO weighted MSE minimization. Section 3 describes the spatial UEP bit-loading based on virtual diagonalization using MSE duality. Section 4 discusses the results. Finally, we conclude our findings in the last section.

II. Multiuser MIMO-OFDM Adaptive Model

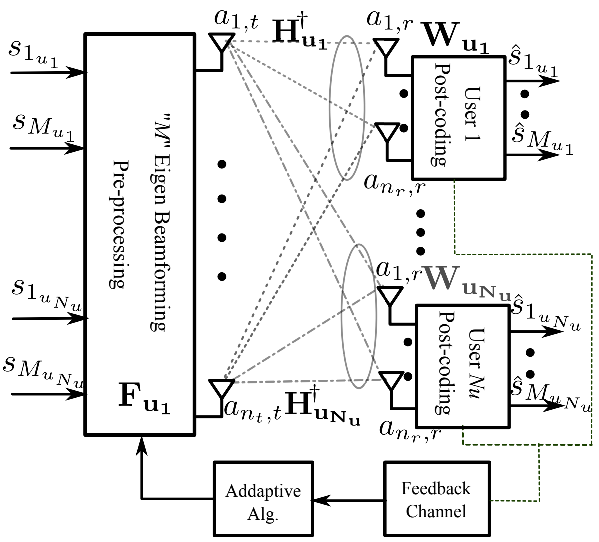

Fig. 1. Adaptive MIMO BC model with the channel $\mathbf{H}^\dagger$

According to the system model depicted in Fig. 1, we consider an $n_r \times n_t$ MIMO BC channel matrix $\mathbf{H}^\dagger_{k,u}$ [1], where n_r is the number of the receive antennas of the u^{th} user for N_u users, n_t is the number of the transmit antennas at the base station, and k is the subcarrier index. The total number of subcarriers is N. We assume a MIMO channel version $\mathbf{H}^\dagger_{k,u} \in \mathbb{C}\mathcal{N}(0, \sigma^2_{\mathbf{H}}\mathbf{I})$ [2], where the channel matrix entries are uncorrelated zero mean circularly symmetric complex Gaussian (ZMCSCG) values and modeled as independent Rayleigh fading blocks with an exponentially decaying power-delay profile [12].

[1] $\mathbf{A}^\dagger$ denote the Hermitian transpose of $\mathbf{A}$

A. *Formulation and Weighted Sum-MSE*

Since our approach uses the MAC-BC duality, one can solve the dual (virtual) MAC scenario to find the transmit and the receive filters. Thereafter, these filters are linearly mapped to the original scenario [6] Thus, for a dual MAC, the transmit and the receive antennas are interchanged, i.e., the mobile terminals act (virtually) as the transmitters, where each mobile terminal is loaded with n_r (virtual transmit) antennas, while the base-station is deploying an n_t (virtual receive) antennas. This is simply done by considering the original channel matrix to be $\mathbf{H}^\dagger_u$, while the virtual MAC remains $\mathbf{H}_u$. Hence, the overall equivalent (dual) uplink received signal at the u^{th} user, after passing through the receiver post-processing $\mathbf{W} \in \mathbb{C}^{n_t \times n_r}$, is given by

$$\mathbf{r} = \mathbf{W} \sum_{i=1}^{N_u} \mathbf{H}_i \mathbf{F}_i \mathbf{\Psi}_i \mathbf{x}_i + \mathbf{W}\mathbf{n} \,, \tag{2}$$

where $\mathbf{W} = [\mathbf{W}_1 \mathbf{W}_2 ... \mathbf{W}_u ... \mathbf{W}_{N_u}]^{\mathrm{T}}$ and $\mathbf{W}_u$ is the dual individual post-processing for each user. $\mathbf{F}_u \in \mathbb{C}^{n_r \times n_r}$ is the user's dual pre-coding matrix, $\mathbf{\Psi}_u \in \mathbb{C}^{n_t \times n_r}$ is a spectral shaping matrix, and $\mathbf{n} \in \mathbb{C}^{n_t \times 1}$ is the AWGN vector with zero mean and variance σ_n^2. In order to find the minimum weighted sum-MSE, the following optimization problem must be solved

$$\begin{aligned} &\text{minimize}_{\mathbf{F},\mathbf{W}} && \sum_{i=1}^{N_u} w_i E_i \\ &\text{subject to} && \sum_{i=1}^{N_u} \mathrm{Tr}\left(\mathbf{F}_i^\dagger \mathbf{F}_i\right) \leq P_m \,, \end{aligned} \tag{3}$$

where E_u denotes the mean-squared error of user u symbols and w_i is a weighting factor required to adapt the QoS. Assuming a total transmit power of $P_m = 1$. Knowing that the receiver filter given in [5], [6] to be a MMSE receiver, which minimizes the MSE for each user individually, we can write

$$\begin{aligned} \mathbf{W}_u &= \mathbf{F}_u^\dagger \mathbf{H}_u^\dagger \mathbf{T}^{-1} \\ &= \mathbf{F}_u \mathbf{H}_u \left[\sigma_n^2 \mathbf{I}_{n_r} + \sum_{k=1}^{N_u} \mathbf{H}_u \mathbf{F}_u \mathbf{F}_u^\dagger \mathbf{H}_u^\dagger \right]^{-1} . \end{aligned} \tag{4}$$

The following algorithm shows how to fulfill the Karush-Kuhn-Tucker (KKT) conditions and achieve the MMSE using the iterative scaled gradient projection methods. The conversion of this algorithm is discussed in [7, Theorem 2].

B. Prioritized Multiuser transmission

The optimum resource allocation based on the weighted sum-MSE in (3) aims at minimizing the sum-MSE multiplied by a weighting vector $\mathbf{w} = [w_1, .., w_{N_u}]^{\mathrm{T}}$, where $\mathbf{1} \cdot \mathbf{w} = 1$. w can be arbitrarily chosen in order to realize UEP among the given users [8]. Thus, achieving this minimum sum-MSE results by switching off users with weaker channel gains and, instead, concentrate the power only on the good users. This is known in adaptive schemes as *the greedy method*, which can easily maintain a QoS-based transmission. If $w_1 = w_2 = ... = w_{N_u} = 1/N_u$, the minimization problem in (3) yields an optimum uncoded symbol-error ratio SER [6].

In order to solve the constrained optimization problems in (3), the authors in [6] proposed to solve a standard unconstrained gradient algorithm which is modified in order to accommodate the power constraint as follows [10], [7]:

$$\mathbf{F}^{(t+1)} = \left[\mathbf{F}^{(t)} - \eta \mathbf{M}^{-1} \nabla E(\mathbf{F}^{(\mathbf{t})})\right]^{+} , \quad (5)$$

where E is the total error given in [7], η is the step size, ∇ corresponds to the matrix-valued Nabla operator (Jacobian matrix) [10, Chapter 3], and the notation $[x]^+$ denotes the orthogonal projection (with respect to the Euclidean norm) of a vector x onto its convex set, say, X. In particular, $[x]^+$ is defined as in [10, Chapter 3] by $[x]^+ = \arg\min_{z \in X} ||z - x||_2$. $\mathbf{M}$ represents a preconditioning positive definite diagonal matrix, which is chosen (according to [10, Section 3.2]) to be

$$\mathbf{M}^{-1}(t) = \sqrt{\frac{P_m}{\sum_{i=1}^{N_u} \nabla E\left(\mathbf{F}_i^{(t)}\right)}} \mathbf{I} , \quad (6)$$

where the denominator in this equation has a very small value. Therefore, this scaling is independent of the SNR. One of the sub-optimality reasons is that this scheme always guarantees similar power allocation across different streams at high SNR, i.e., the scaling is extremely small. $\nabla E_u \left(\mathbf{F}^{(t)}\right)$ is defined (also similar to the Lagrangian method in [11]) as

$$\begin{aligned} \nabla E_u \left(\mathbf{F}_u^{(t)}\right) &= \frac{\partial E_u(\mathbf{F_u}^{(t)})}{\partial \mathbf{F}_u^{\dagger(t)}} \\ &= -\mathbf{H}_u^{\dagger} \mathbf{T}^{-1} \left(w_i \mathbf{T} - S\right) \mathbf{T}^{-1} \mathbf{H}_u \mathbf{F}_u \, . \end{aligned} \quad (7)$$

Now, substitute (7) into (6) then into (5) to achieve an accurate $\mathbf{F}_u$. Once $\mathbf{F}_u$ is computed, $\mathbf{T}$ is calculated using $\mathbf{H}_u$ and $\mathbf{F}_u$ using (4). Finally, $\mathbf{W}_u$ is computed using (4); further iterations are required for enhancing the results.

The details of these steps are implemented in [6, Algorithm 1] using an iterative gradient projection method. This algorithm fulfills the KKT conditions to achieve the MMSE for all users. The conversion of this algorithm is also discussed in [7, Theorem 2].

C. Uplink/downlink conversion:

After the transceiver filters are computed for the dual problem, these filters have to be converted to the original scenario. The conversion to the original filters is performed based on the following steps:

1) compute the scaling factor α_0 for the first user as in [6, Algorithm 1, line 12],

$$\alpha_0 = \sqrt{\frac{P_m}{\mathrm{Tr}\left(\sum_{i=1}^{N_u} w_i \mathbf{P}_i^{\dagger} \mathbf{H}_i^{\dagger} \mathbf{T}^{-2} \mathbf{H}_i \mathbf{P}_i\right)}} , \quad (8)$$

where α is the factor used to model the near-far effect in the dual MAC.
2) compute the scaling factor for the remaining users using their weighting values such that: $\alpha_u = \alpha_0 \sqrt{w_u}$
3) find the downlink transceiver filters

$$\begin{aligned} \mathbf{F}^{\mathrm{DL}} &= [\alpha_u \mathbf{T}^{-1} \mathbf{H}_u \mathbf{F}_1 ... \alpha_u \mathbf{T}^{-1} \mathbf{H}_u \mathbf{F}_{N_u}]^{\mathrm{T}} \\ \mathbf{W}_u^{\mathrm{DL}} &= \frac{1}{\alpha_u} \mathbf{F}_u^{\dagger} \alpha \end{aligned} \quad (9)$$

III. Virtual Diagonalization and Throughput Maximization

A. Per-user Sum-rate

The previous MSE minimization does not guarantee independence (diagonalization) across users, which is the main goal in realizing adaptive –and prioritized– transmission. Even more, individual user streams are not jointly optimized and, hence, cannot be used directly by any rate-adaptive algorithm. Therefore, it is required to minimize the MSE of each stream separately. However, the question here will be: can we make these streams independent? On the one hand, this is a very complex optimization approach. On the other hand, if even complexity is not an issue, it will not be optimum solution from the capacity point of view. The reason for sub-optimality is the non-convexity of (3) in $\{\mathbf{F}_u, \mathbf{W}_u\}$ [3], which results into local minima. Therefore, it is required to imitate, e.g., the block-diagonalization. This virtual diagonalization can only be achieved if the SINR is used in the sum-rate formula such that

$$R_u^i = \log_2\left(1 + \mathrm{SINR}_u\right) \quad \forall\, i \in \{1..M_u\} \, , \quad (10)$$

Therefore, one can exploit the relation between the SINR and the MMSE. This relation is found to be, similar to the linear MMSE equalizer derivation given in [4, Chapter 10] or the bijective mapping [6]:

$$\mathrm{MMSE}_u = \frac{1}{1+\mathrm{SINR}_u}\,, \tag{11}$$

where MMSE_u is the summation of all individual MMSE for the data streams used by user u. Hence, the per-user sum-rate is

$$R_u = -\log_2\left(\mathrm{MMSE}_u\right)\,. \tag{12}$$

This can still give a good estimate for the overall data rate for the u^{th} user's streams without the need for optimizing each stream separately [7].

B. UEP Bit-loading

Similar, to the algorithm in [12], the $\mathrm{MMSE}_{u,k}$ for each user u and for every subcarrier k are sorted in descending order according to their MMSE values (from low to high). Thus, bits of the first class are allocated to the stronger subcarriers (with the lowest MMSE). Accordingly, the noise margins of the users γ_u are iteratively adapted to fulfill the required bit-rate using the same approximation in [9], which requires few iterations to converge. However, an accurate noise margin γ_u needs to be computed first. Therefore, the average number of bits allocated to all data streams is

$$b_{u,k} \approx -\log_2\left(\mathrm{MMSE}_{u,k}\gamma_u\right)\,. \tag{13}$$

C. Power allocation

The power is automatically allocated via the gradient projection method pre-coding matrix $\mathbf{F}_u$. However, the power allocated from this algorithm suffers from sub-optimality due to allocating similar power values across the streams and subcarriers at high-SNR values (as mentioned in Section II). Therefore, we implemented a spectral shaping matrix $\mathbf{\Psi}$ which, first, devotes zeros to the location where the water-filling would allocate zero-power. Second, after the bit-loading is calculated using the algorithm in [12], the power on the non-zero subcarriers are allocated according to bit load values and the user margin γ_u.

IV. Results

Figure 2 depicts the performance of the non-adaptive prioritized weighted sum-MSE. In this case, $w_0 > w_1 > w_2$, i.e., user_0 has the highest priority, while user_2 has the least one. These results have been generated using 2 bits/symbol/stream, i.e., 4 bits/user. In the middle of these curves (thick blue) one can see the performance of the non-weighted sum-MSE minimization, as an average of the highest and the lowest priority user.

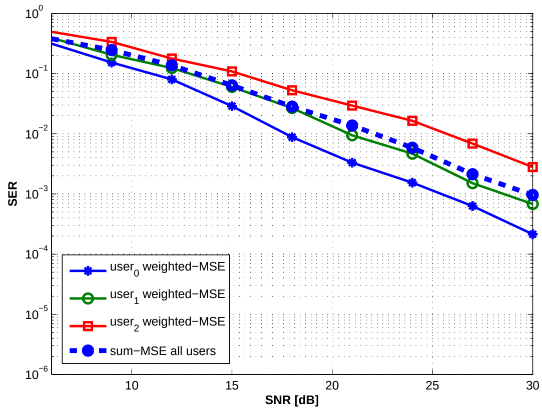

Fig. 2. Sum-MSE vs Weighted MSE with 4/bits/user

In Fig. 4, the performance of the three users with weighted sum-MSE are depicted using UEP bit-loading with a margin separation between UEP classes of 3 dB. From this figure, one can notice that the double prioritized transmission is achieved with (almost) equal separations between classes and users. Moreover, in Fig. 4, the adaptive UEP bit-loading is depicted for the first and the third users using equivalent 2 bits/symbol/stream. One can easily notice a performance gain of 12 dB (between the first user's highest protected class and the same user using non-adaptive modulation). However, at low-SNR subcarriers, the non-adaptive scheme outperforms the adaptive one due to overloading the high-SNR subcarriers. In Fig. 5, the UEP bit-loading performance (using a non-weighted sum-MSE) is considered to deliver an average performance between the highest and the least protected users.

V. Conclusions

We described a double-prioritized adaptive UEP transmission. First, a weighted MMSE, utilizing the MAC-BC duality, has been implemented, which realizes different QoSs. Second, a UEP adaptive modulation that allows for arbitrary margin separations combined with the prioritized weighted sum-MSE was considered. This means to allow for different user QoS among users and data. Results in case of adaptive (using virtual UEP bit-loading) multiuser MIMO-OFDM outperforms the non-adaptive case by almost 12 dB. This is due to utilizing

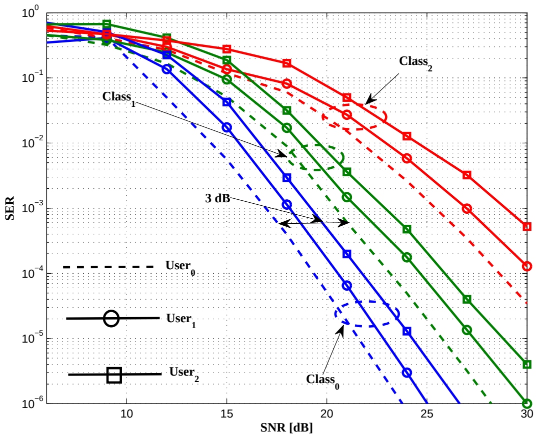

Fig. 3. SER for MIMO-OFDM multiuser with 3 users, each with 3 data priority classes with 1365 bits/class

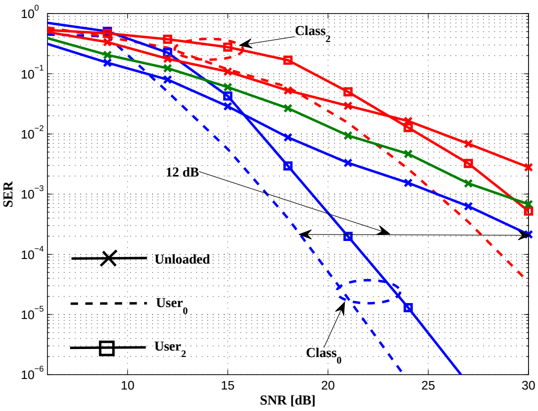

Fig. 4. SER for adaptive and non-adaptive modulation

the subcarriers with the lower MMSE. The non-prioritized non-weighted sum-MSE remains at an average performance between users with prioritized transmission.

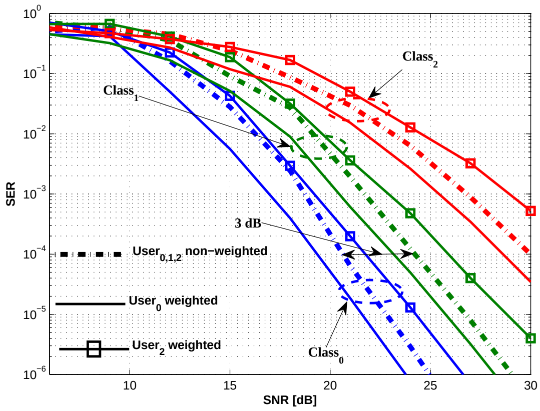

Fig. 5. Sum-MSE vs Weighted MSE using 4096 bits/1024 subcarriers

ACKNOWLEDGMENT

This work is funded by the German National Science Foundation (Deutsche Forschungsgemeinschaft, DFG).

REFERENCES

[1] **Bourdoux, A. and Khaled, N.**, "Joint TX-RX Optimization for MIMO-SDMA Based on a Null-Space Constraint," proc. *Vehicular Technology Conference, VTC 2002* Birmingham, UK, 6-9 May 2002, vol.1, pp. 171-174.

[2] **Hassan, K. and Henkel, W.**, "Prioritized Multi-layer Transmission for MIMO-OFDM Multiuser Broadcast Channel," *14th International OFDM-Workshop 2009*, Hamburg, Germany, Sept. 2 - 3, 2009.

[3] **Schubert, M. and Boche, H.**, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *Vehicular Technology, IEEE Transactions on, 2004*, Jan, vol. 53, pp. 18 - 28.

[4] **Proakis, J.G.**, *Digital Communications*, 4^{th} Ed., McGraw-Hill, New York, 2001.

[5] **Shuying Shi, Schubert, M., and Boche, H.**, "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: MMSE Balancing," *Signal Processing, IEEE Transactions on*, vol. 56, no. 8, pp. 3702–3712, Aug. 2008.

[6] **Hunger, R., Joham, M., and Utschick, W.**, "On the MSE-Duality of the Broadcast Channel and the Multiple Access Channel," *Signal Processing, IEEE Transactions on, 2009*, vol. 57, pp. 698 - 713.

[7] **Mezghani, A., Joham, M., Hunger, R., and Utschick, W.**, "Transceiver designs for multiuser mimo systems," *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 698–713, Feb. 2009.

[8] **Shuying Shi, Schubert, M., and Boche, H.**, "Rate Optimization for Multiuser MIMO Systems With Linear Processing," *Signal Processing, IEEE Transactions on, 2008*, vol. 56, pp. 4020 - 4030.

[9] **Hassan, K., Sidhu, G., and Henkel, W.**, "Multiuser MIMO-OFDMA with Different QoS Using a Prioritized Channel Adaptive Technique," *ICC Workshops 2009*, Dresden, Germany, 2009, pp. 1 - 5.

[10] **Bertsekas, D.P., and Tsitsiklis, J.N.**, "Parallel and distributed computation," *Englewood Cliffs, NJ*, 1999.

[11] **Serbetli, S. and Yener, A.**, "Transceiver optimization for multiuser MIMO systems," *Signal Processing, IEEE Transactions on*, vol. 52, no. 1, pp. 214–226, Jan. 2004.

[12] **Hassan, K. and Henkel, W.**, "UEP MIMO-OFDM with Beamforming-Combining for Imperfect Channel Information," *12th International OFDM-Workshop 2007*, Hamburg, Germany, Aug. 27 - 28, 2007.