Toward URLLC: A Full-Duplex Relay System with Self-Interference Utilization or Cancellation

Yufei Jiang, Hanjun Duan, Xu Zhu, Zhongxiang Wei, Tong Wang, Fu-Chun Zheng, and Sumei Sun

Abstract—**Ultra-reliable and low-latency communication (URLLC) is one of the key use cases of the fifth generation (5G) wireless communications to facilitate specific application scenarios with stringent latency and reliability demands, such as industrial automation and Tactile Internet. A full-duplex (FD) relay with simultaneous transmission and reception in the same frequency band is an effective approach to enhance the reliability of cell-edge user terminals, by significantly suppressing self-interference (SI). However, the signal processing latency at FD relay due to SI cancellation, referred to as relaying latency, takes a significant part in the end-to-end latency, and therefore should be minimized, while guaranteeing high reliability. In this article, we first present an up-to-date overview of the end-to-end latency for an FD relay system, addressed on physical layer challenges. We investigate the possible solutions in the literature to achieve the goal of URLLC. The efficient solution is to allow a simple amplify-and-forward (AF) FD relay mode with low-complexity SI radio frequency and analog cancellations, and process the residual SI alongside the desired signal at base station in an adaptive manner, rather than being cancelled at relay in digital domain. Also, the residual SI can be utilized at base station to enhance the reliability and the degree of freedom in signal processing, not necessarily being cancelled as much as possible. The FD relay assisted system with adaptive SI utilization or cancellation enables extended network coverage, enhanced reliability and reduced latency, compared to the existing overview work.**

INTRODUCTION

LOW latency and high reliability have become two major challenges in the future wireless communication design. However, more and more emerging devices and applications request strict constraints of high reliability and low latency, such as industrial automation, Tactile Internet, virtual reality, Internet of vehicles, telemedicine and so on [1], as shown in Fig. 1. In the fifth generation (5G) cellular networks, ultra-reliable and low-latency communication (URLLC) is introduced to achieve reliability of 99.999 percent and latency of 1 ms. The future networks are expected to require even more stringent reliability and latency (e.g., below 1 ms).

It is very challenging to achieve such two ambitious targets for cell-edge user terminals that encounter wireless deep fading, dramatically reducing the link reliability. The application of a relay [2], as an efficient way to mitigate wireless fading, can be utilized to guarantee high reliability of a communication link highly separated by a cell-edge user and base station. However, the latency is very high for a traditional half-duplex (HD) relay system, as two time slots are required to complete a single data transmission. The source transmits data to an HD relay in the first time slot, and the HD relay forwards them to destination in the second time slot. Meanwhile, the source is not allowed to transmit data in the second time slot. Thus, the latency of the HD relay system is twice, as much as possible, that of direct transmissions with no relay and no retransmission during data transmission completion. In order to reduce the latency, a full-duplex (FD) relay, equipped with transmit antennas and receive antennas, has been widely studied in the literature, to enable simultaneous transmission and reception in the same frequency band, with theoretically doubled throughput [3]. The source and relay can successively transmit data in consecutive time slots. Thus, the FD relay system is a promising solution to provide latency lower than the HD relay system, and close to direct transmissions with no relay and no retransmission. In the literature, there is only one overview work for an HD relay-enabled URLLC system in the finite blocklength regime [2]. So far, there has been no overview work for an FD relay system from the URLLC perspective.

In an FD relay system, the reliability is limited by self-interference (SI) due to the signal leakage from transmit antennas to receive antennas at relay, seriously affecting bit error rate (BER) performance [4]. Recent breakthroughs reveal up to 120 dB SI cancellation capability [5], and facilitate the real application of FD communications at relay rather than at base station. The SI suppression is as low as approximately -97 dBm at relay with transmission power of 23 dBm, and -74 dBm at base station with transmission power of 46 dBm. The significant SI suppression at FD relay can be as low as the noise power level of -90 dBm. To guarantee high reliability, traditional works [3-6] require the FD relay to work in the decode-and-forward (DF) mode with a number of SI cancellation processes, such as radio frequency (RF) cancellation, analog cancellation and digital cancellation. However, the SI cancellation increases processing latency, referred to as relaying latency, which should be maintained at a low level as much as possible, without compromising reliability in terms of SI cancellation capability, to achieve URLLC.

To the best of the authors' knowledge, this is the first work to present an insightful investigation of reliability and latency together for an FD relay assisted URLLC system. We provide an overview of the end-to-end latency, where the relaying latency plays a significant part. This has not been presented in the existing overview works [2, 5, 7]. We discuss possible relaying latency reduction solutions. An efficient solution is to

Yufei Jiang, Hanjun Duan, Xu Zhu, Tong Wang and Fu-Chun Zheng are with the Harbin Institute of Technology (Shenzhen);

Xu Zhu is also with the University of Liverpool;

Zhongxiang Wei is with the University College London;

Sumei Sun is with the Institute for Infocomm Research, Singapore.

Yufei Jiang and Hanjun Duan equally contributed to this work. (Corresponding authors: Tong Wang and Xu Zhu)

allow the FD relay to work in the amplify-and-forward (AF) mode with low-complexity RF cancellation and analog can cellation, keeping the FD relay at a low-cost manner. This is different from the existing overview works [2, 5], where the DF mode is required with all of three SI cancellation processes, yielding high relaying latency. We also investigate the residual SI cancellation and utilization conducted at base station in digital domain, not necessarily being cancelled as much as possible at FD relay as in [5, 6]. The residual SI can be utilized to improve reliability and enhance the degree of freedom in signal processing, which has not been introduced in the recent overview literatures [2, 5].

We evaluate the end-to-end latency and reliability in terms of BER. The FD relay system with the AF mode provides latency significantly lower than that with the DF mode and the HD relay system, respectively, and close to direct transmissions with no relay and no retransmission. Also, the AF FD relay system with SI utilization at base station provides better BER performance, compared to the DF FD relay system with all SI cancellations at FD relay.

FD RELAY IN URLLC

For FD relay systems, as shown in Fig. 2, the overall latency is addressed on physical layer, and can generally be divided into transmission latency, propagation latency, pre-processing latency, processing latency and relaying latency.

Transmission latency. This corresponds to the time duration between the beginning of a packet transmission and the end of the same packet transmission at the transmitter. The packet design is a key issue to minimize the transmission latency in URLLC. In 5G New Radio systems, a non-square packet stretched in frequency is used to reduce transmission latency [7]. As shown in Fig. 2, there are two parts of transmission latency between source to relay and between relay to destination for relay-based systems. The transmission latency can be minimized at an FD relay with simultaneous reception and transmission. Thus, the FD relay system provides transmission latency close to the direct transmission system with no relay and no retransmission.

Propagation latency. We consider an urban scenario, where the transmission distance is about 500 meters between base station and cell-edge users [4]. There are two courses of signal propagation between source to relay and between relay to destination. Given a user terminal far from the serving base station, as shown in Fig. 3, it is important to select a proper relay to secure two signal propagations in the line-of-sight scenario with strong channel gains and short propagation distances, achieving simultaneous high reliability and low latency.

Pre-processing latency. This corresponds to signaling feedback and exchange at base station, including a number of feedback information such as channel state information (CSI), quality of service requirement, bandwidth and capacity requirements. Also, the control signaling is incorporated, such as hybrid automatic repeat request (HARQ), connection request message, scheduling grant message, queuing latency and so on.

Processing latency. This part includes optimization latency and signal processing latency. In URLLC, short-frame transmissions are allowed, with a number of resources available, such as frequency, bandwidth, channel, time and power resources, the number of relays, the number of antennas at relay and at base station, the number of symbols in a data frame, queue state information and packet loss probability. With the given resources, a low-complexity cross-layer optimization scheme is preferable to achieve low-latency and high-reliability requirements, formulated with tractable solutions in a closed form. It has been shown in [1] that a cross-layer optimization problem is formulated to minimize transmission power in a closed form under the required quality of service, subject to a number of constraints on physical and media access control layers, such as packet dropping and joint power, bandwidth and subcarrier allocations. In signal processing, it is not available to employ traditional long-frame-based channel coding, such as turbo codes with long redundancy check, which reduces data rates and increases latency in finite block length. It is possible to use control-signaling-based channel coding in URLLC, for example, polar codes in control channels in 5G [8]. Advanced channel coding is worthwhile investigating for FD relay in URLLC to be robust against the SI. The other signal processing parts are related to channel estimation, carrier frequency offset (CFO) estimation, I/Q imbalance compensation, phase noise estimation, requiring a number of pilots. In the literature, lots of estimation approaches have been proposed with pros and cons. However, the number of pilots is limited in short-frame transmissions. Therefore, the same pilots can be utilized efficiently as much as possible to accomplish multi-task. It has been shown in [9] that a single pilot is used to jointly estimate SI, CFO and channel for FD systems, providing high reliability and low latency.

Relaying latency. Compared with direct transmissions, there is an extra latency, referred to as relaying latency, including RF SI cancellation, analog SI cancellation and digital SI cancellation, which should be maintained at a low level for FD relay systems. **RF SI cancellation** is performed via antenna shielding and isolation plus directional or dual-polarized antennas. The SI can be suppressed by around 45 dB, providing a neglectable latency in propagation domain due to very short signal propagation between receive and transmit antennas at relay. If there is a strong line-of-sight component of SI, **analog SI cancellation** must be operated to cancel the significant component before the received signals transfer into a power amplifier. There are two reasons:

- First, the received power from SI is much higher than that from the source. Thus, the signal received from the source is hardly decoded correctly, if analog cancellation is not used.
- Second, the received signal from the source and SI could provide power higher than the limit of a power amplifier. The signal distortion occurs when the received signal reaches the saturation area of a power amplifier.

The residual SI is the non-line-of-sight component that can be reduced in digital domain. Traditionally, **digital SI**

cancellation allows the relay to work in the DF mode, requiring a complex signal processing based on an equivalent baseband model in a chip with a complex design, which increases processing latency at relay, unable to meet the low-latency requirement. An efficient solution is to allow digital SI cancellation at base station rather than at relay, reducing relaying latency. The base station is equipped with high-quality electronic components and chips, and provides high computation capability in digital SI cancellation with low signal processing latency. The details are described in next sections.

DIGITAL CANCELLATION

Assume that orthogonal frequency division multiplexing (OFDM) modulation is employed. There is an integer OFDM symbol processing latency at FD relay. The latency is at least one OFDM symbol to ensure that the received symbol is not correlated with the transmitted symbol within one OFDM symbol at FD relay [6]. Traditionally, as shown in Fig. 4 (a), digital SI cancellation is conducted at FD relay to mitigate the residual SI signals from the output of analog cancellation. The CSI of SI is estimated from transmit antennas to receive antennas at FD relay by a number of pilots. The SI signal is regenerated using the estimated channel and the decoded signal in the previous OFDM symbol. The regenerated SI signal is subtracted from the received signal in the current OFDM symbol to obtain a clear signal with no SI before transmission from the transmit antenna at FD relay.

In order to achieve the goal of URLLC, digital SI cancellation can be processed at base station rather than at relay [10-12]. The main reasons can be listed as follows:

- **Low relaying latency**: Compared to direct transmissions, relaying latency takes a significant part in the end-to-end latency for the FD relay system, and should be minimized, by allowing FD relay to work in the AF mode with no complex decoding and encoding process. Complex digital SI cancellation can be pushed as much as possible to base station with high calculation capability to achieve low latency.
- **Low-cost hardware of relay**: Generally, a base station is surrounded by a number of relays to extend the coverage range and improve the reliability of cell-edge users. Digital SI cancellation that works at a number of DF FD relays requires multiple sets of complex chips and hardwares, while just one set is probably required at base station to cancel residual SI digitally with the AF FD relay used. Therefore, the AF FD relay can reduce the cost of hardware, and a large number of low-cost AF FD relays can be deployed.

However, there is insufficient work to investigate digital SI cancellation at base station.

SI power attenuation. The previous OFDM symbol from the transmit antenna at relay is taken as an SI interfering with the current OFDM symbol from cell-edge users, which forms an SI loop with cumulative effect over all previous OFDM symbols. In other words, the SI consists of summing up all previous OFDM symbols, if the residual SI is not cancelled at relay. However, the power of the residual SI consecutively fades exponentially with the number of transmitted OFDM symbols in the past, as the residual SI channel gain is relatively low [10-12], with the line-of-sight component being mitigated by analog cancellation. Therefore, the current OFDM symbol is only interfered with a small number of previously successive OFDM symbols [10-12] that take a significant part in the SI power. The residual SI power of the other left OFDM symbols is too low, and can be treated as noise [10-12].

Digital cancellation procedure. Digital SI cancellation is preferred to perform in frequency domain based on each subcarrier, allowing a number of element-wise calculations in SI regeneration instead of high-complexity pseudo-inverse and convolution operations in time domain at base station. As shown in Fig. 4 (b), the digital cancellation procedure at base station is described as follows:

- SI includes equivalent channels from transmit antenna to receive antenna at relay, and further to base station. There is no need to estimate channels separately, as the equivalent channels can be estimated by pilots transmitted together with source data. The received signal includes equivalent channels from cell-edge users to relay, and further to base station. The equivalent channels can also be estimated by pilots inserted in data.
- SI signal is reconstructed via a number of previously decoded OFDM symbols and SI channels.
- SI is subtracted digitally from the current received OFDM symbol at base station.

To some extent, the channel estimation error destroys the reconstruction of SI signal, and gives rise to additional interference and system degradation. Also, the number of OFDM symbols considered as SI depends on the signal to residual SI power ratio (SIR). It has been shown in [11] that a number of two previous OFDM symbols dominate 99.9 percent power of residual SI when the ratio of the SI to relay-to-destination channel variances is as low as -40 dB.

Precoding-based SI Cancellation. For a downlink, digital SI cancellation is not preferable at a user terminal, since most of mobile phones and devices are composed of cheap electronic components and are incapable of dealing with complex signal processing. Hence, it is preferable to perform precoding-based SI cancellation at a base station with high computation capability, allowing low relaying latency. The reconstructed SI signal with the known CSI is subtracted from the desired signal before transmission, which does not dramatically reduce the signal power, because the desired signal provides channel gain higher than SI signal. The propagation of the desired signal between base station and relay is line-of-sight, while the SI channels are non-line-of-sight components after analog cancellation at relay. The SI can be cancelled automatically at the receive antennas of relay, when the received signal is superimposed with the SI from the transmit antennas of relay.

RESIDUAL SI UTILIZATION

As SI originates from previous OFDM symbols, the residual SI can be utilized to enhance reliability and the degree of freedom in signal processing rather than being cancelled as much as possible in digital domain, when being transmitted together with the desired signal to base station.

Virtual MISO systems. The residual SI is a delayed version of the desired signal, and can be modelled as a virtual multiple-input single-output (MISO) system [11] together with the desired signal, if there is a single receive antenna at FD relay and at base station, respectively. Thus, the residual SI can be treated as a self-coding to build a structure of space time code at the received signals at base station to enhance the degree of freedom in signal processing. It has been shown in [12] that a spatial diversity is built at base station with two virtual antennas, where one received signal is from the relay and the other one is from the cell-edge user via direct transmissions. The desired signal is mixed with the SI signal, to improve reliability in terms of BER, based on bit-interleaved coded modulation [12].

Virtual MIMO systems. The residual SI signal and the desired signal can be viewed as multiple input signals to build a virtual multiple-input multiple-output (MIMO) system, if multiple antennas are equipped at transmit and receive sides of FD relay as well as at base station. As shown in Fig. 4 (c), a receive and a transmit antennas are grouped into a transmission link at relay. The received signal at AF FD relay is first operated by analog SI cancellation, and then re-transmitted to multi-antenna base station. By building an equivalent MIMO channel model, the desired and SI signals are equalized on each subcarrier in frequency domain at base station. The equalized SI is a replica of previous OFDM symbols, and is utilized to combine with the equalized signal in previous OFDM symbols to maximize the desired signal power for enhancement of degree of freedom in signal processing as well as reliability improvement. It has been shown that up to 10 dB signal power enhancement can be yielded [10]. The procedure of SI utilization is described as follows:

- Build an equivalent virtual MIMO system model with respect to the SI and desired signals.
- Estimate SI and the desired signal channels.
- Employ zero-forcing or minimum mean squared error criterion to equalize the desired and SI signals.
- Apply the equalized SI signals to enhance the power of the desired signal in previous OFDM symbols.

In order to successfully equalize the desired and SI signals, the number of antennas required is the same as or larger than the number of OFDM symbols taken as SI plus one desired signal. For example, if the SI power is dominated by the previous one OFDM symbol, the SI signal generated from the previous OFDM symbol and the desired signal are used to form two virtual inputs, which requires at least two-receive and two-transmit antennas at FD relay along with at least two-receive antennas at base station. The degree of freedom in signal processing is enhanced at expense of a number of additional antennas used. This is not a big problem, as current and future devices are equipped with multiple antennas.

Blind source separation. Since the desired and SI signals can be formulated as a linear MIMO model, a number of blind source separation approaches [9, 10] can be employed to separate the received signals blindly, such as subspace [9] and independent component analysis [10]. Blind source separation provides high spectral efficiency to recover the desired and SI signals with no knowledge of CSI and therefore no pilot required, which can reduce the impact of channel estimation errors resulting from a limited number of pilots used in short-frame transmission in URLLC. It has been shown in [10] that the previous one OFDM symbol is considered as a significant component in the SI and can be extracted by blind source separation, with the residual OFDM symbols in the SI being treated as noise, when SIR is at least around 20 dB. However, the separated signals present some ambiguity drawbacks in terms of phase and permutation. Phase ambiguity is referred to a number of phase shifts in the equalized signals, while permutation ambiguity is to sort out the disorder problem in the separated signals. In other words, the blindly separated signals should be recognized which is the desired signal, and which is the SI signal. The ambiguities can be eliminated by precoding or short pilots. Nevertheless, blind source separation is based on data statistics, and requires data frame that is not too short.

Millimeter-wave transmissions. FD relay is applied for millimeter-wave transmissions to extend network coverage, as the propagation attenuation is significantly strong in small wavelength. Due to the fact that millimeter-wave transmissions are easily blocked by obstacles, SI power can be significantly reduced by employing antenna shielding between receive and transmit sides of relay and highly directional transmit antennas with narrow beamwidth directly to the destination [13]. The SI channels between receive and transmit antennas of relay can be modelled as Rayleigh fading collected from reflected waves, with strong line-of-sight path being blocked. Thus, analog cancellation is not required at relay, and the recursive loopback residual SI can be cancelled in digital domain at base station if the relay works in the AF mode. Also, the residual SI can be utilized to enhance the degree of freedom in signal processing at base station with multiple antennas.

RELIABILITY AND LATENCY EVALUATION

Reliability evaluation. Figure 5 shows BER performance of the AF FD relay system with digital SI cancellation or utilization at base station. The AF FD relay system with SI utilization at base station provides BER lower than 10^{-5}, from signal to noise ratio (SNR) being 22 dB to 30 dB, which meets the 99.999 percent reliability requirement of URLLC [7], equivalent to 10^{-5} in terms of BER [10]. Also, its reliability improves by about 4 dB, compared to the DF FD relay system with all SI cancellations at relay. The AF FD relay with digital SI cancellation at base station is shown to provide BER performance close to the DF FD relay system.

Latency evaluation. In Table I, the latency evaluation for the AF FD relay is shown, in comparison to the DF HD relay, the DF FD relay and direct transmissions with no relay and no retransmission. There are three significant latency components: transmission latency, relaying latency and processing latency. The transmission latency is the time-to-transmit, corresponding to an OFDM symbol of 17.86 μs in URLLC [7]. Due to simultaneous transmission and reception at FD relay, the FD relay system provides transmission latency

TABLE I
EVALUATION OF LATENCY
SYSTEM LEVEL PARAMETERS ARE SET AS FOLLOWS. THE SUBCARRIER SPACING IS 60 KHZ. THE BANDWIDTH IS SET AS 20 MHZ. RF AND ANALOG CANCELLATIONS ARE UP TO 30 DB AND 40 DB, RESPECTIVELY. TRANSMISSION POWER RANGES FROM 10 DBM TO 30 DBM, AND NOISE FLOOR IS -95 DBM (DC: DIGITAL CANCELLATION, BS: BASE STATION).

Transmission category	Accomplished steps		Average time consumed (μs)	Latency of each OFDM symbol (μs)
DF HD relay [14]	Transmission latency [7]		35.72	99.609
	Relaying latency	Channel estimation [14]	25.669	
		Signal detection [14]	36.565	
	Processing latency at BS (with high computation capability)	Channel estimation	1.655	
		Signal detection		
DF FD relay [3]	Transmission latency [7]		17.86	46.028
	Relaying latency	Analog SI cancellation [3]	0.122	
		SI channel estimation [3]	26.359	
		SI signal reconstruction and mitigation [3]	0.032	
	Processing latency at BS (with high computation capability)	Channel estimation	1.655	
		Signal detection		
AF FD relay with DC at BS	Transmission latency [7]		17.86	20.551
	Relaying latency	Analog SI cancellation	0.122	
	Processing latency at BS (with high computation capability)	Channel estimation	2.569	
		SI signal reconstruction and mitigation		
		Signal detection		
AF FD relay with SI utilization at BS [10]	Transmission latency [7]		17.86	19.747
	Relaying latency	Analog SI cancellation	0.122	
	Processing latency at BS (with high computation capability)	Equalization and signal separation	1.765	
		SI utilization		
Direct transmissions with no relay and no retransmission	Transmission latency [7]		17.86	19.515
	Processing latency at BS (with high computation capability)	Channel estimation	1.655	
		Signal detection		

close to direct transmissions, and lower than the HD relay system that requires one time slot for data transmission and the other time slot for data reception. Thus, the transmission latency of the FD relay system is approximately as much as that of direct transmissions, and a half of that of the HD relay system, as shown in Table I. The relaying latency and the processing latency are evaluated by simulation, and are determined by the computational complexity of the signal processing algorithm, such as channel estimation algorithm, signal detection algorithm and SI cancellation method, requiring a number of matrix multiplications and matrix inversions. The computational complexity is represented as the latency evaluation, characterized by the time running on CPU. In simulation, the associated calculations are simulated to run over 10,000 OFDM symbols to obtain the average time consumption on each OFDM symbol for each algorithm [15]. Also, the relaying latency and the processing latency are affected by the baseband chip. The baseband chip at base station generally provides computation capability higher than that at relay. In URLLC, the retransmission is allowed in the HARQ protocol when block error rate is lower than 10^{-3}, and the maximum number of retransmissions is 2 [7]. Hence, when cell-edge users are far from base station, the AF FD relay system with SI cancellation or utilization at base station can guarantee higher reliability with lower block error rate, requiring less retransmissions, compared to direct transmissions in deep fading. The queuing latency in URLLC can be modelled as statistical queuing requirement, characterized by the maximum queuing latency of 0.8 ms and a small latency violation probability [1]. It is demonstrated in Table I that the AF FD relay system with digital SI cancellation or utilization at base station provides the overall latency of an OFDM symbol close to the direct transmission with no relay and no retransmission, and significantly lower than the DF HD relay system and the DF FD relay system, respectively.

CONCLUSION

In this article, the reliability and latency of an FD relay have been overviewed for URLLC-enabled systems. The relaying latency is a significant part in the end-to-end latency, and can be minimized, while guaranteeing high reliability. To reduce the relaying latency, the FD relay is discussed to work in the AF mode with low-complexity RF SI cancellation and analog SI cancellation. The residual SI can be processed together with the desired signals at base station in digital domain. Also, the residual SI can be utilized to improve the reliability and enhance the degree of freedom in signal processing. The FD relay system with the AF mode provides latency significantly lower than that with the DF mode and the HD relay system, respectively, and close to direct transmissions with no relay and no retransmission. Also, the AF FD relay system with SI utilization at base station provides better BER performance, compared to the DF FD relay system with all SI cancellations at FD relay.

ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China under Grants 61901138 and 61801145, in part by Natural Science Foundation of Guangdong Province under Grants 2018A030313298 and

2018A030313344, in part by Shenzhen Science and Technology Program under Grants JCYJ20180306171800589 and KQTD20190929172545139, and in part by Guangdong Science and Technology Planning Project under Grant 2018B030322004.

REFERENCES

[1] C. She, C. Yang, and T. Q. S. Quek, "Cross-layer optimization for ultra-reliable and low-latency radio access networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, Jan. 2018, pp. 127–141.

[2] Y. Hu, M. C. Gursoy, and A. Schmeink, "Relaying-enabled ultra-reliable low-latency communications in 5G," *IEEE Network*, vol. 32, no. 2, Apr. 2018, pp. 62–68.

[3] M. Mohammadkhani Razlighi and N. Zlatanov, "Buffer-aided relaying for the two-hop full-duplex relay channel with self-interference," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, Jan. 2018, pp. 477–491.

[4] B. Ma, H. Shah-Mansouri, and V. W. S. Wong, "Full-duplex relaying for D2D communication in millimeter wave-based 5G networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, July 2018, pp. 4417–4431.

[5] M. S. Sim *et al.*, "Nonlinear self-interference cancellation for full-duplex radios: From link-level and system-level performance perspectives," *IEEE Communications Magazine*, vol. 55, no. 9, Sept. 2017, pp. 158–167.

[6] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, Sept. 2011, pp. 3074–3085.

[7] H. Ji *et al.*, "Ultra-reliable and low-latency communications in 5G downlink: Physical layer aspects," *IEEE Wireless Communications*, vol. 25, no. 3, June 2018, pp. 124–130.

[8] A. Jalali and Z. Ding, "Joint detection and decoding of polar coded 5G control channels," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, Jan. 2020, pp. 2066–2078.

[9] Y. Liu *et al.*, "Fast iterative semi-blind receiver for URLLC in short-frame full-duplex systems with CFO," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 4, Apr. 2019, pp. 839–853.

[10] H. Duan *et al.*, "An adaptive self-interference cancelation/utilization and ICA-assisted semi-blind full-duplex relay system for LLHR IoT," *IEEE Internet of Things Journal*, vol. 7, no. 3, Mar. 2020, pp. 2263–2276.

[11] Y. Liu, X. Xia, and H. Zhang, "Distributed linear convolutional space-time coding for two-relay full-duplex asynchronous cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, Dec. 2013, pp. 6406–6417.

[12] Y. Jin *et al.*, "Full-duplex delay diversity relay transmission using bit-interleaved coded OFDM," *IEEE Transactions on Communications*, vol. 65, no. 8, Aug. 2017, pp. 3250–3258.

[13] Z. Wei *et al.*, "Full-duplex versus half-duplex amplify-and-forward relaying: Which is more energy efficient in 60-GHz dual-hop indoor wireless systems?" *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 12, Dec. 2015, pp. 2936–2947.

[14] W. Liu, C. Li, and J. Li, "Achieving maximum degrees of freedom of two-hop MIMO alternate half-duplex relaying system for linear transceivers: A unified transmission framework for DF and AF protocols," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 2144–2148, July 2015.

[15] X. Quan *et al.*, "Blind nonlinear self-interference cancellation for wireless full-duplex transceivers," *IEEE Access*, vol. 6, July 2018, pp. 37725–37737.

BIOGRAPHIES

Yufei Jiang [S'12, M'14] (Jiangyufei@hit.edu.cn) received his Ph.D. degree in Electrical Engineering and Electronics from the University of Liverpool, Liverpool, U.K., in 2014. From 2014 to 2015, he was a Postdoctoral Researcher with the Department of Electrical Engineering and Electronics, University of Liverpool. From 2015 to 2017, he was a Research Associate with the Institute for Digital Communications, University of Edinburgh, Edinburgh, U.K. He is currently an Assistant Professor with the Harbin Institute of Technology, Shenzhen, China. His research interests include Li-Fi, synchronization, full-duplex, and blind source separation.

Hanjun Duan [S] (duanhanjun@stu.hit.edu.cn) received a B.S. degree in Electronics and Information Engineering from the College of Applied Science of Jiangxi University of Science and Technology, Ganzhou, China, in 2013, and an M.S. degrees in Communication and Information Systems from the Yunnan Minzu University, Kunming, China, in 2016. He is currently working toward the Ph.D. degree at the School of Electrical and Information Engineering, Harbin Institute of Technology, Shenzhen, China. His research interests include channel estimation and equalization, full-duplex, and blind source separation.

Xu Zhu [S'02, M'03, SM'12] (xuzhu@liverpool.ac.uk) received a B.Eng. degree from Huazhong University of Science and Technology, Wuhan, China, in 1999, and a PhD degree from Hong Kong University of Science and Technology, Hong Kong, in 2003. She is a Reader of University of Liverpool, Liverpool, UK, and also with Harbin Institute of Technology, Shenzhen, China. She has more than 190 peer-reviewed publications. Her research interests include MIMO, channel estimation and equalization, resource allocation, green communication etc. She has served as Editor for IEEE Transactions on Wireless Communications and Symposium Co-Chair of IEEE ICC 2016, ICC 2019 and Globecom 2021.

Zhongxiang Wei [S'15, M'17] (zhongxiang.wei@ucl.ac.uk) received his Ph.D. degree in electrical and electronics engineering from the University of Liverpool, United Kingdom, in 2017. From March 2016 to March 2017, he was with the Institution for Infocomm Research, Agency for Science, Technology, and Research (A*STAR), Singapore, as a research assistant. He is currently a research associate in electrical and electronics engineering at University College London, United Kingdom. His research interests include constructive interference design, green communications, full-duplex, millimeter-wave communications, and algorithm design.

Tong Wang [M] (tongwang@hit.edu.cn) received a B.Eng. degree in electrical engineering and automation from Beihang University, Beijing, China, in 2006 and an M.Sc. degree in communications engineering and a Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2008 and 2012, respectively. From 2012 to 2015, he was a Research Associate in RWTH Aachen University, Aachen, Germany. From 2014 to 2015, he was a Research Fellow of the Alexander von Humboldt Foundation. Since March 2016, he has been with the School of Electronics and Information Engineering, Harbin Institute of Technology, Shenzhen, China, where he is an Assistant Professor.

Fu-Chun Zheng [M'95, SM'99] (zhengfuchun@hit.edu.cn) was with the University of Reading, UK, from September 2007 to July 2016 as a Professor (Chair) of Signal Processing. Since August 2016, he has been with Harbin Institute of Technology (Shenzhen), China, as a distinguished professor. His current research interests include multiple antenna systems, URLLC, green communications, and ultra-dense networks. He served as the general chair of IEEE VTC 2006-S, Melbourne, Australia (the first ever VTC held in the southern hemisphere). More recently he was the executive TPC Chair for VTC 2016-S, Nanjing, China (the first ever VTC held in mainland China).

Sumei Sun [F'16] (sunsm@i2r.a-star.edu.sg) is a Principal Scientist and Head of Communications and Networks Dept at the Institute for Infocomm Research (I^2R), Singapore. She is also holding a joint appointment with Singapore Institute of Technology, and an adjunct appointment with National University of Singapore, both as a full professor. She is Editor-in-Chief of IEEE Open Journal of Vehicular Technology, member of IEEE Transactions on Wireless Communications Steering Committee. She's also Director of IEEE Communications Society Asia Pacific Board, Chapter Coordinator of Asia Pacific Region in IEEE Vehicular Technologies Society, and member of IEEE Communications Society Globecom/ICC Management and Strategy Standing Committee.

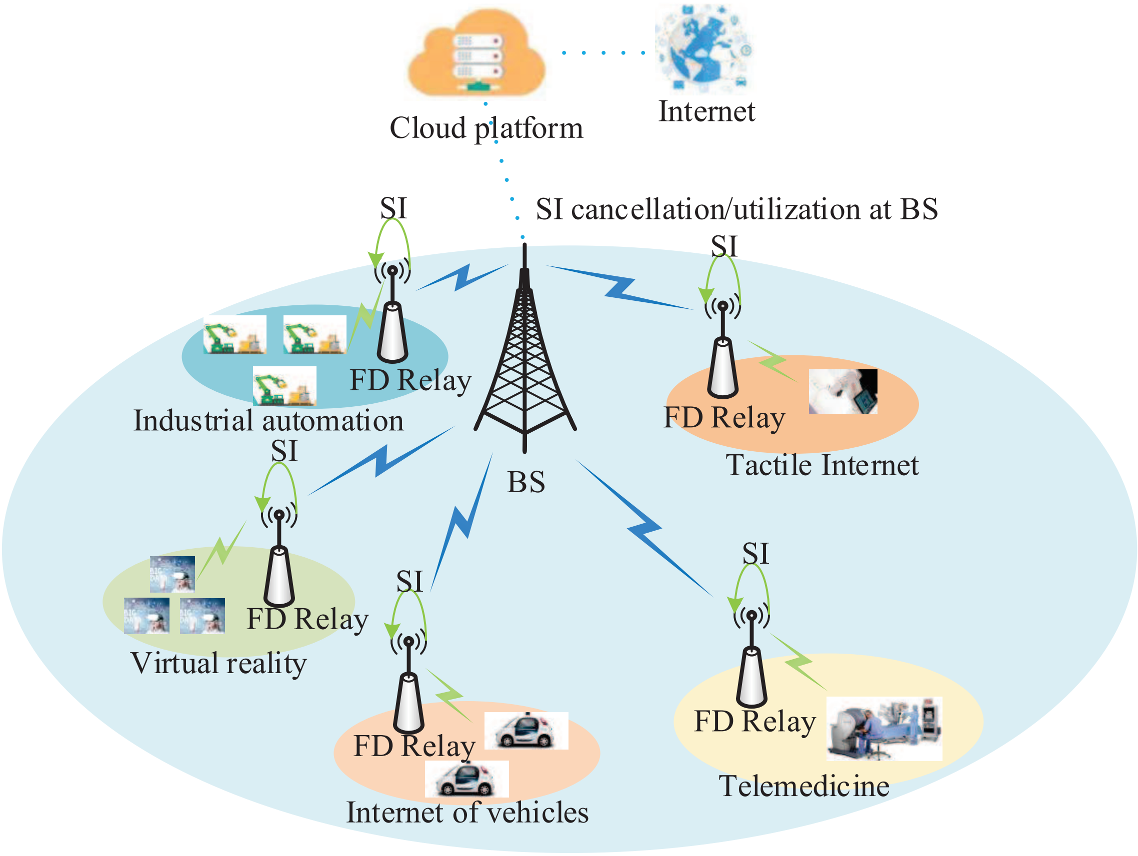


Fig. 1. Emerging devices and applications that require ultra-reliable and low-latency demands. User terminals far from base station are assisted by an FD relay to guarantee high reliability and low latency, allowing an FD relay to work in the AF mode and pushing digital SI cancellation/utilization at base station to reduce relaying latency (BS: base station).

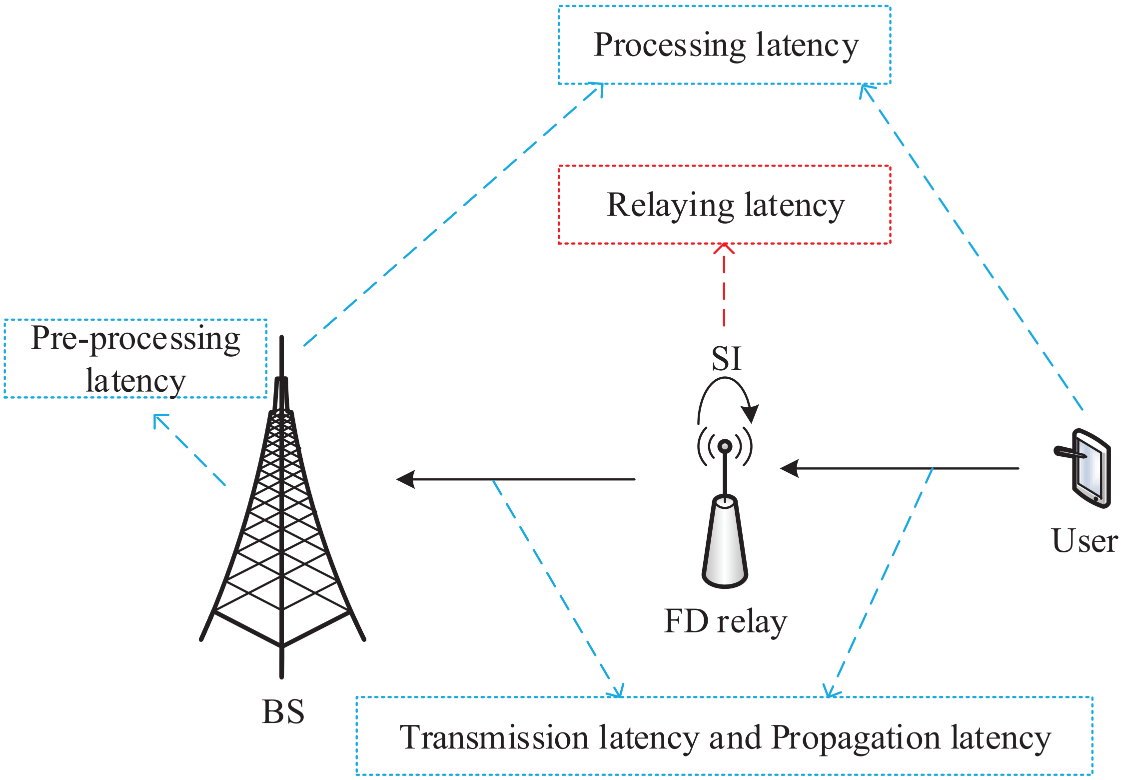


Fig. 2. Latency diagram for FD relay systems in physical layer with main components: pre-processing latency, relaying latency, processing latency, transmission latency and propagation latency (BS: base station).

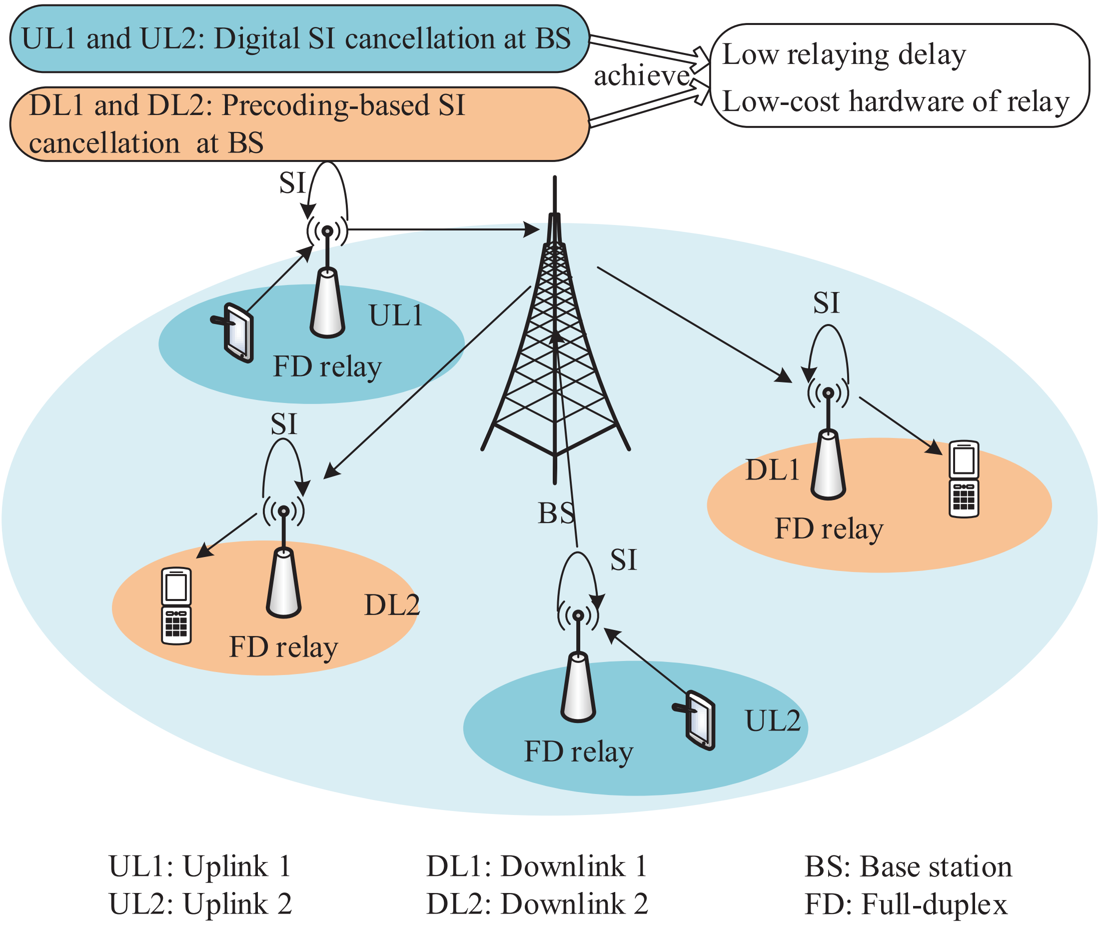
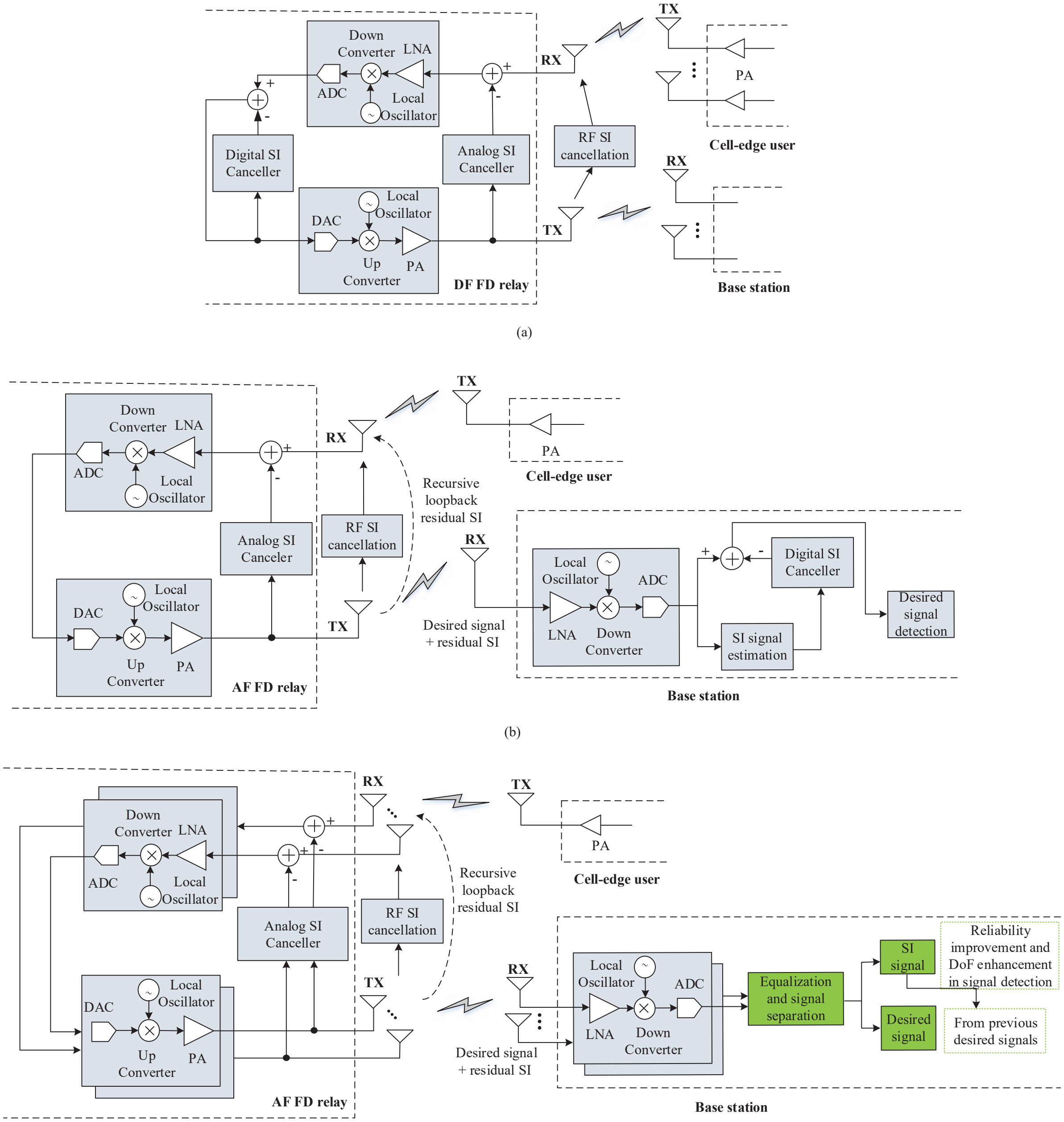


Fig. 3. Diagram of residual SI cancellation and utilization in digital domain at base station with uplink in blue circle and downlink in orange circle as well as relay selection.



(a)

(b)

(c)

Fig. 4. (a) For traditional DF FD relay systems, the SI is cancelled at relay, using RF, analog and digital cancellations, which leads to an increase in relaying latency; (b) An AF FD relay system model is presented, with a single transmit and a single receive antennas at relay as well as a single antenna at base station. In order to meet low latency requirement, analog and RF cancellations are necessary to perform at relay. The residual SI is transmitted together with the desired signal to base station for further process. (c) An AF FD relay system model is presented, with multiple transmit and multiple receive antennas at relay as well as multi-antenna base station. The residual SI after analog and RF cancellations is transmitted together with the desired signal to base station in digital domain, and is utilized to improve reliability and enhance the degree of freedom in signal processing. (LNA: low-noise amplifier, PA: power amplifier, ADC: analog-to-digital converter, DAC: digital-to-analog converter, RX: receiver, TX: transmitter, DoF: degree of freedom).

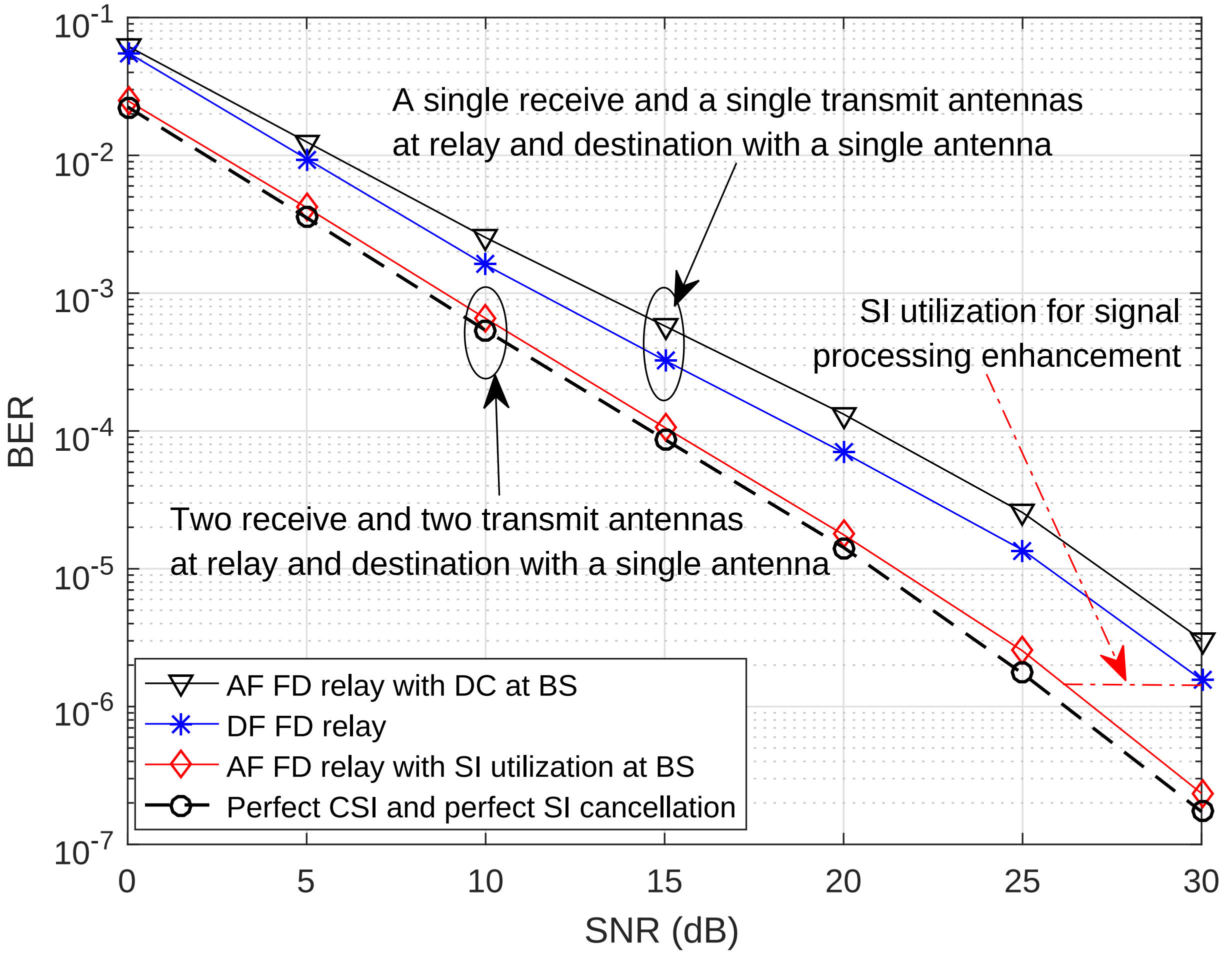


Fig. 5. BER performance of the AF FD relay system with digital SI cancellation or utilization at base station, in comparison to the traditional DF FD relay system with all SI cancellations at relay. A slot level is considered with 7 OFDM symbols [7]. The subcarrier spacing is 60 kHz. The bandwidth is set as 20 MHz. RF and analog cancellations are up to 30 dB and 40 dB, respectively. Transmission power ranges from 10 dBm to 30 dBm, and noise floor is -95 dBm. SIR is 20 dB (DC: digital cancellation).